

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein -- protein search, using sw model

Run on: June 5, 2003, 16:12:50 ; Search time 27 seconds
(without alignments)
545.959 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLMAGVLPAAH.....CLRLRQHQHDFADDISLKK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2656	99.7	501	4	US-09-548-372D-4
2	2656	99.7	501	4	US-09-548-367D-4
3	2650	99.5	501	4	US-09-009-191-2
4	2582	96.9	501	4	US-09-713-158-2
5	2567	96.4	501	4	US-09-548-372D-8
6	2567	96.4	501	4	US-09-548-372D-8
7	2506.5	94.1	476	4	US-09-548-372D-6
8	2506.5	94.1	476	4	US-09-548-372D-6
9	2420.5	90.9	476	4	US-09-548-372D-73
10	2420.5	90.9	476	4	US-09-548-372D-73
11	2397	90.0	453	4	US-09-548-372D-30
12	2397	90.0	453	4	US-09-548-372D-30
13	2397	90.0	459	4	US-09-548-372D-32
14	2397	90.0	459	4	US-09-548-372D-32
15	2315	86.9	774	4	US-09-009-191-4
16	2291.5	86.0	446	4	US-09-548-372D-22
17	2291.5	86.0	446	4	US-09-548-372D-22
18	2288	85.9	433	4	US-09-548-372D-26
19	2288	85.9	433	4	US-09-548-372D-26
20	2288	85.9	459	4	US-09-548-372D-24
21	2288	85.9	459	4	US-09-548-372D-24
22	2247.5	84.4	428	4	US-09-548-372D-51
23	2247.5	84.4	428	4	US-09-548-372D-51
24	2247.5	84.4	434	4	US-09-548-372D-53
25	2247.5	84.4	434	4	US-09-548-372D-53
26	2104	79.0	425	4	US-09-548-372D-28
27	2104	79.0	425	4	US-09-548-372D-28

28	1173.5	44.1	518	3	US-08-999-723-2	Sequence 2, Appli
29	1173.5	44.1	518	4	US-09-434-427-2	Sequence 2, Appli
30	1173.5	44.1	518	4	US-09-548-372D-2	Sequence 2, Appli
31	1173.5	44.1	518	4	US-09-548-367D-2	Sequence 2, Appli
32	1150	43.2	514	4	US-09-717-432-2	Sequence 2, Appli
33	1150	43.2	514	4	US-09-912-484-2	Sequence 2, Appli
34	301.5	11.3	396	1	US-08-208-007A-13	Sequence 13, Appl
35	301.5	11.3	396	4	US-09-032-523-9	Sequence 9, Appli
36	301.5	11.3	396	4	US-08-915-095A-13	Sequence 13, Appl
37	301.5	11.3	396	4	US-08-798-096-13	Sequence 13, Appl
38	301.5	11.3	396	4	US-08-798-095A-13	Sequence 13, Appl
39	300.5	11.3	412	1	US-08-208-007A-12	Sequence 12, Appl
40	300.5	11.3	412	4	US-08-974-691-4	Sequence 4, Appli
41	300.5	11.3	412	4	US-08-915-095A-12	Sequence 12, Appl
42	300.5	11.3	412	4	US-08-798-096-12	Sequence 12, Appl
43	300.5	11.3	412	4	US-08-798-095A-12	Sequence 12, Appl
44	282.5	10.6	458	6	5217891-15	Patent No. 5217891
45	276.5	10.4	427	2	US-08-846-021A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-4

Query Match	99.7%	Score 2656;	DB 4;	Length 501;
Best Local Similarity	99.8%	Pred. No. 2.9e-267;	Mismatches 1;	Indels 0; Gaps 0;
Matches 500;	Conservative 0;			
Qy	1	MAQALPWLMLMAGVLPAAHGTQHGI	RLPLRSLGGLGAPLGLRLPRETDEEPEE	RRGSGF 60
Db	1	MAQALPWLMLMAGVLPAAHGTQHGI	RLPLRSLGGLGAPLGLRLPRETDEEPEE	RRGSGF 60
Qy	61	VEWVDNLGRSGGGYVEMTVGSPPTL	NILVDTGSSNFVAGAAPHPFLHRYQ	ORLSST 120
Db	61	VEWVDNLGRSGGGYVEMTVGSPPTL	NILVDTGSSNFVAGAAPHPFLHRYQ	ORLSST 120
Qy	121	YRDLRKGVVYPYTGCKWEGELGTDL	SVTPHGPNTVVRANTAAITESDKFF	INGSNWEGIL 180
Db	121	YRDLRKGVVYPYTGCKWEGELGTDL	SVTPHGPNTVVRANTAAITESDKFF	INGSNWEGIL 180
Qy	181	GLAYAEIARPDSDLPEPFDSLVKQ	THVPLNPLFSLHLCGAPPLNQSEVL	ASVGGSMIIIGI 240
Db	181	GLAYAEIARPDSDLPEPFDSLVKQ	THVPLNPLFSLHLCGAPPLNQSEVL	ASVGGSMIIIGI 240
Qy	241	DHSLYTGSWYTPIRREMYVEVIVR	VEINGQDLKMDCKEYNYDKSIVDS	GTTNLRPKK 300
Db	241	DHSLYTGSWYTPIRREMYVEVIVR	VEINGQDLKMDCKEYNYDKSIVDS	GTTNLRPKK 300

Db 121 YRDLRKGYYEPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360
Db 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Db 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Qy 481 RCLRCRQOHHDFADDISLLK 501
Db 481 RCLRCRQOHHDFADDISLLK 501

RESULT 4

US-09-713-158-2
; Sequence 2, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: POWELL, DAVID J.
; APPLICANT: CHRISTIE, GARY
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)
; FILE REFERENCE: GP-70660
; CURRENT APPLICATION NUMBER: US/09/713.158
; PRIOR FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/165,800
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
9-713-158-2

Query Match 96.9%; Score 2582; DB 4; Length 501;
Best Local Similarity 96.6%; Pred. No. 1.4e-259;
Matches 484; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
Qy 1 MAQALPMLLLMWGACVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Db 1 MAQALPMLLLMWGACVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Qy 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDGTSSNFVAVCAAPHPHRLHYRQQLSST 120
Db 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDGTSSNFVAVCAAPHPHRLHYRQQLSST 120
Qy 121 YRDLRKGYYEPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYYEPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360
Db 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Db 421 HVHDEFRTAAVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVQW 480
Qy 481 RCLRCRQOHHDFADDISLLK 501
Db 481 RCLRCRQOHHDFADDISLLK 501

RESULT 5

US-09-548-372D-8
; Sequence 8, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-8

Query Match 96.4%; Score 2567; DB 4; Length 501;
Best Local Similarity 96.2%; Pred. No. 5.1e-258;
Matches 482; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MAQALPMLLLMWGACVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Db 1 MAQALPMLLLMWGACVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60
Qy 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDGTSSNFVAVCAAPHPHRLHYRQQLSST 120
Db 61 VEMVDNLKRGSGQGYVEMTVGSPPTNLILVDGTSSNFVAVCAAPHPHRLHYRQQLSST 120
Qy 121 YRDLRKGYYEPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYYEPTQGWEGELGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVFNLSHLGCGAGFPLNOSSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Db 241 DHSLYTGLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360
Db 301 VFEAAVKSIIKASSTKEKFPDGFNLGEQVLCVQWAGTTPWNIIPVLSLYLMGEVTTNOSFRIT 360

QY 361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420
DB 361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
QY 481 RCLRLRQOHHDDFADDISLLK 501
DB 481 RCLRLRQOHHDDFADDISLLK 501
RESULT 6
US-09-348-367D-8
; Sequence 8, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-348-367D-8
Query Match 96.4%; Score 2567; DB 4; Length 501;
Best Local Similarity 96.2%; Pred. No. 5.1e-258;
Matches 482; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
QY 1 MAQALPWLMLLMWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGGSF 60
DB 1 MAPALHLLMLLMWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGGSF 60
61 VEMVDNLKRGKSGQGYVYVEMTSGSPPTNLILVDTGSSNFVAGAAPHPFLHRYYQRLSST 120
61 VEMVDNLKRGKSGQGYVYVEMTSGSPPTNLILVDTGSSNFVAGAAPHPFLHRYYQRLSST 120
121 YRDLRKGVPVYPTQKWEGLGTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
121 YRDLRKGVPVYPTQKWEGLGTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
181 GLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
301 VFEAAVKSIIKAASSTKFPDGFGLGQVLCWAGTTPWNIFPVISLYLMGEVNTQSPRIT 360
301 VFEAAVKSIIKAASSTKFPDGFGLGQVLCWAGTTPWNIFPVISLYLMGEVNTQSPRIT 360
361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420
DB 361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420

QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
QY 481 RCLRLRQOHHDDFADDISLLK 501
DB 481 RCLRLRQOHHDDFADDISLLK 501
RESULT 7
US-09-548-372D-6
; Sequence 6, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-6
Query Match 94.1%; Score 2506.5; DB 4; Length 476;
Best Local Similarity 95.0%; Pred. No. 9.1e-252;
Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
QY 1 MAQALPWLMLLMWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGGSF 60
DB 1 MAQALPWLMLLMWAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEPEEPPGRGGSF 60
61 VEMVDNLKRGKSGQGYVYVEMTSGSPPTNLILVDTGSSNFVAGAAPHPFLHRYYQRLSST 120
61 VEMVDNLKRGKSGQGYVYVEMTSGSPPTNLILVDTGSSNFVAGAAPHPFLHRYYQRLSST 120
121 YRDLRKGVPVYPTQKWEGLGTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
121 YRDLRKGVPVYPTQKWEGLGTDLVSIHPGPNVTVRANIAAITESDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
181 GLAYAEIARPDSDLPEFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTGTTNLRPKK 300
301 VFEAAVKSIIKAASSTKFPDGFGLGQVLCWAGTTPWNIFPVISLYLMGEVNTQSPRIT 360
301 VFEAAVKSIIKAASSTKFPDGFGLGQVLCWAGTTPWNIFPVISLYLMGEVNTQSPRIT 360
361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420
DB 361 ILPQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVYVDFRARRKIGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVQW 480
QY 481 RCLRLRQOHHDDFADDISLLK 501

Db 456 RCLRCLRQHQHDFADDISLLK 476

RESULT 8

US-09-548-367D-6
; Sequence 6, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548.367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-6

Query Match 94.1%; Score 2506.5; DB 4; Length 476;
Best Local Similarity 95.0%; Pred. No. 9.1e-252;
Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

Qy 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60
Db 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60

Qy 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNLIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNLIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120

Qy 121 YRDLKRGVVPYTGKWEGLGTLVSIPLHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLKRGVVPYTGKWEGLGTLVSIPLHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

Qy 181 GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
Db 181 GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 215

Qy 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 300
Db 216 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 275

Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVVCWAGTTPWNIPFVLSLYLMGEVTVNQSPRIT 360
Db 276 VFEAAVKSIIKAASSTKFPDGFGLGEOLVVCWAGTTPWNIPFVLSLYLMGEVTVNQSPRIT 335

Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 420
Db 336 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 395

Qy 421 HVHDEFRTAAGVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 480
Db 396 HVHDEFRTAAGVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 455

Qy 481 RCLRCLRQHQHDFADDISLLK 501
Db 456 RCLRCLRQHQHDFADDISLLK 476

RESULT 9

US-09-548-372D-73
; Sequence 73, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548.372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 73
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-548-372D-73

Query Match 90.9%; Score 2420.5; DB 4; Length 476;
Best Local Similarity 91.8%; Pred. No. 8e-243;
Matches 460; Conservative 5; Mismatches 11; Indels 25; Gaps 1;

Qy 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60
Db 1 MAQALPMLLLMMGAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60

Qy 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNLIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNLIVDTGSSNFVAGCAAPHPFLHRYQRLSST 120

Qy 121 YRDLKRGVVPYTGKWEGLGTLVSIPLHGPNTVVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLKRGVVPYTGKWEGLGTLVSIPLHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

Qy 181 GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 240
Db 181 GLAYAEIARPPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMITGGI 215

Qy 241 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 300
Db 216 DHSLYTGLWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDSTTNLRLPKK 275

Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVVCWAGTTPWNIPFVLSLYLMGEVTVNQSPRIT 360
Db 276 VFEAAVKSIIKAASSTKFPDGFGLGEOLVVCWAGTTPWNIPFVLSLYLMGEVTVNQSPRIT 335

Qy 361 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 420
Db 336 ILPOQYLRPVEDVATSDQDCYKFAISQSTGTVMGAVIMEGYVYVDFRARRKRGFAVSAC 395

Qy 421 HVHDEFRTAAGVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 480
Db 396 HVHDEFRTAAGVEGFPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCQW 455

Qy 481 RCLRCLRQHQHDFADDISLLK 501
Db 456 RCLRCLRQHQHDFADDISLLK 476

RESULT 10

US-09-548-367D-73
; Sequence 73, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 73

; LENGTH: 476

; TYPE: PRT

; ORGANISM: Mus musculus

; US-09-548-367D-73

Query Match

Best Local Similarity 90.9%; Score 2420.5; DB 4; Length 476;

Mismatches 5; Indels 25; Gaps 1;

Matches 460; Conservative 5;

QY 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPCRRGSF 60

DB 1 MAPALHLLWVSGMLPAQHTLGLRLPLRSLGGLGAPLGLRLPRETDEEPEEPCRRGSF 60

QY 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDGTSSNFAYGAAPHPLHRYQRLSST 120

DB 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDGTSSNFAYGAAPHPLHRYQRLSST 120

QY 121 YRLRGVVPYPTQGWEGELGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

DB 121 YRLRGVVPYPTQGWEGELGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

QY 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240

DB 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240

QY 241 DLSLYTGSLSWYTPIRREWYEVIIIRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

DB 241 DLSLYTGSLSWYTPIRREWYEVIIIRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

QY 216 DLSLYTGSLSWYTPIRREWYEVIIIRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 275

QY 301 VFEAAVKSIRAASTKTEKFPDGLGQVLCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360

DB 276 VFEAAVKSIRAASTKTEKFPDGLGQVLCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 335

QY 361 ILPQOYLRPVEDVATSDODCKYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420

DB 336 ILPQOYLRPVEDVATSDODCKYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 395

QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES 453

DB 396 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES 476

QY 481 RCLRLRQHQHDDFADDSILK 501

DB 456 RCLRLRQHQHDDFADDSILK 476

RESULT 11

US-09-548-372D-30

Sequence 30, Application US/09548372D

Patent No. 6420534

GENERAL INFORMATION:

APPLICANT: GURNEY ET AL.

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 29915/6280I

CURRENT APPLICATION NUMBER: US/09/548,372D

CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 30

; LENGTH: 453

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-548-372D-30

Query Match

Best Local Similarity 99.8%; Pred. No. 2e-240;

Mismatches 452; Conservative 0; Indels 0; Gaps 0;

QY 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPCRRGSF 60

DB 1 MAQALPWLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPCRRGSF 60

QY 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDGTSSNFAYGAAPHPLHRYQRLSST 120

DB 61 VEMVDNLKSGGGYVEMTVGSPPTNLILVDGTSSNFAYGAAPHPLHRYQRLSST 120

QY 121 YRLRGVVPYPTQGWEGELGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

DB 121 YRLRGVVPYPTQGWEGELGTDLVSIPIHGPNTVVRANIAAITESDKFFINGSNWEGIL 180

QY 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240

DB 181 GLAYAEIARPDSDLEPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240

QY 241 DLSLYTGSLSWYTPIRREWYEVIIIRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

DB 241 DLSLYTGSLSWYTPIRREWYEVIIIRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

QY 301 VFEAAVKSIRAASTKTEKFPDGLGQVLCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360

DB 301 VFEAAVKSIRAASTKTEKFPDGLGQVLCWQAGTTPWNIFFVISLYLMGEVNTNQSFRIT 360

QY 361 ILPQOYLRPVEDVATSDODCKYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420

DB 361 ILPQOYLRPVEDVATSDODCKYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420

QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES 453

DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDES 453

RESULT 12

US-09-548-367D-30

Sequence 30, Application US/09548367D

Patent No. 6440698

GENERAL INFORMATION:

APPLICANT: GURNEY ET AL.

TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

TITLE OF INVENTION: THEREOF

FILE REFERENCE: 29915/6280H

CURRENT APPLICATION NUMBER: US/09/548,367D

CURRENT FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: US 60/155,493

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: US 09/404,133

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: PCT/US99/20881

PRIOR FILING DATE: 1999-09-23

PRIOR APPLICATION NUMBER: US 60/101,594

PRIOR FILING DATE: 1998-09-24

NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 453
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-548-367D-30

Query Match 90.0%; Score 2397; DB 4; Length 453;
Best Local Similarity 99.8%; Pred. No. 2e-240;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Qy 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGGFPNLSQSEVLAASVGGSMIIGGI 240
Db 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGGFPNLSQSEVLAASVGGSMIIGGI 240
Qy 241 DHSLYTGLWYTPIRREMYEYIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRLPKK 300
Db 241 DHSLYTGLWYTPIRREMYEYIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRLPKK 300
Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIIPVLSLYLMGEVTTNQSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIIPVLSLYLMGEVTTNQSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDCCYKFAISQSSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDCCYKFAISQSSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Qy 421 HVHDEFRATAAEGFPVTLDMEDCGYNIPQTDDES 453
Db 421 HVHDEFRATAAEGFPVTLDMEDCGYNIPQTDDES 453

RESULT 13
US-09-548-372D-32
; Sequence 32, Application US/09548372D
; Patent No. 6420334
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-548-372D-32

Query Match 90.0%; Score 2397; DB 4; Length 459;
Best Local Similarity 99.8%; Pred. No. 2.le-240;

Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Qy 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Db 121 YRDLRKGYVYPYTOGKWEGLGTLVSIPIHGNVTVRANIAAITESDKFFINGSNWEGIL 180
Qy 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGGFPNLSQSEVLAASVGGSMIIGGI 240
Db 181 GLAYAEATARPDDSLPEPFDLSLVKQTHVFNLSHLGAGGFPNLSQSEVLAASVGGSMIIGGI 240
Qy 241 DHSLYTGLWYTPIRREMYEYIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRLPKK 300
Db 241 DHSLYTGLWYTPIRREMYEYIIVRVEINGQDLKMDCKEYNDKSIIVDSCTTNLRLPKK 300
Qy 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIIPVLSLYLMGEVTTNQSFRIT 360
Db 301 VFEAAVKSIIKAASSTKFPDGFGLGEQLVCWQAGTTPWNIIPVLSLYLMGEVTTNQSFRIT 360
Qy 361 ILPOQYLRPVEDVATSDCCYKFAISQSSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Db 361 ILPOQYLRPVEDVATSDCCYKFAISQSSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
Qy 421 HVHDEFRATAAEGFPVTLDMEDCGYNIPQTDDES 453
Db 421 HVHDEFRATAAEGFPVTLDMEDCGYNIPQTDDES 453

RESULT 14
US-09-548-367D-32
; Sequence 32, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 459
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-548-367D-32

Query Match 90.0%; Score 2397; DB 4; Length 459;
Best Local Similarity 99.8%; Pred. No. 2.le-240;
Matches 452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLWAGAGVLPAGHGTQHGIRLPLRSLGGGAPLGLRLPRETDEEPEEGRGGSF 60
Qy 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120
Db 61 VEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFVAGAAAPHPLHRYQRLSST 120

[illegible]

RESULT 15

US-09-009-191-4
Sequence 4, Application US/09009191
Patent No. 6319689
GENERAL INFORMATION:
APPLICANT: POWELL, DAVID
APPLICANT: CHAPMAN, CONRAD
APPLICANT: MURPHY, KAY
APPLICANT: SMITH, TRUDI
TITLE OF INVENTION: ASP2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,191
FILING DATE: 20-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9701684.4
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70368
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-009-191-4

```

Query Match      86.9%; Score 2315; DB 4; Length 774;
Best Local Similarity 98.6%; Pred. No. 1.7e-231;
Matches 438; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      58  GSFEVMDNLRGKSGGYVEMTVGSPPTLNLVDTGSSNFVGAAPHFLHRYQRL 117
Db      1  GSFEVMDNLRGKSGGYVEMTVGSPPTLNLVDTGSSNFVGAAPHFLHRYQRL 60
      |||
QY      118 SSTYRDLRKGYVPTQGWKEGELGTDLVSIPHGPNTVVRANIAAITESDKFFINGSWE 177
Db      61 SSTYRDLRKGYEPTQGWKEGELGTDLVSIPHGPNTVVRANIAAITESDKFFINGSWE 120
      |||
QY      178 GILGLAYAEIARPDSDLPEPFDLSLVKQTHVPNLFSLHLCGAGFPLNQSEVLASVGGSMII 237
Db     121 GILGLAYAEIARPDSDLPEPFDLSLVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMII 180
      |||
QY      238 GGIDHSLYTGSGLWTPILRRWEYVEYIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRL 297
Db     181 GGIDHSLYTGSGLWTPILRRWEYVEYIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRL 240
      |||
QY      298 PKKFEAAVKSIIKAASSTERKFPDGFWMLEQVLCWQAGTTPWNIFFPVISLYLMGEVTVNSF 357
Db     241 PKKFEAAVKSIIKAASPREKFPDGFWMLEQVLCWQAGTTPWNIFFPVISLYLMGEVTVNSF 300
      |||
QY      358 RTITLPOOYLRPVEDVATSDDCYKFAISQSSSTGTVMGAVIMEGFFVYVDFDRARKRIGFV 417
Db     301 RTITLPOOYLRPVEDVATSDDCYKFAISQSSSTGTVMGAVIMEGFFVYVDFDRARKRIGFV 360
      |||
QY      418 SACHVHDEFRTAAVEGPPVTLDMDCCGYNIPTQDDESTLMTIAYVMAAICALFMLPLCLMV 477
Db     361 SACHVHDEFRTAAVEGPPVTLDMDCCGYNIPTQDDESTLMTIAYVMAAICALFMLPLCLMV 420
      |||
QY      478 CQWRCLRLRQOQDDFADDISLLK 501
Db     421 CQWRCLRLRQTMDDFADDISLLK 444
      |||

Search completed: June 5, 2003, 16:18:45
Job time : 29 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:12:20 ; Search time 24 Seconds
(without alignments)
2155.141 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLLMGAGVLPAAH.....CLRLCROOHDFADDISLLK 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

1 number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2564	100.0	501	9	US-09-869-414-4
2	2664	100.0	501	10	US-09-794-927-4
3	2664	100.0	501	10	US-09-795-847-4
4	2664	100.0	501	10	US-09-794-743-4
5	2664	100.0	501	10	US-09-794-748-4
6	2664	100.0	501	10	US-09-794-925-4
7	2664	100.0	501	10	US-09-681-442-4
8	2656	99.7	501	9	US-10-032-818-4
9	2656	99.7	501	9	US-10-214-932-104
10	2582	96.9	501	9	US-09-969-671A-2
11	2582	96.9	488	9	US-09-795-903A-2
12	2582	96.9	488	9	US-10-032-818-2
13	2582	96.9	488	10	US-09-796-264-2
14	2582	96.9	488	10	US-09-845-226-2
15	2582	96.9	503	9	US-09-795-903A-3
16	2582	96.9	503	9	US-10-032-818-3
17	2582	96.9	503	10	US-09-796-264-3
18	2582	96.9	503	10	US-09-845-226-3
19	2567	96.4	501	9	US-09-869-414-8

20	2567	96.4	501	10	US-09-794-927-8	Sequence 8, Appli
21	2567	96.4	501	10	US-09-795-847-8	Sequence 8, Appli
22	2567	96.4	501	10	US-09-794-743-8	Sequence 8, Appli
23	2567	96.4	501	10	US-09-794-748-8	Sequence 8, Appli
24	2567	96.4	501	10	US-09-794-925-8	Sequence 8, Appli
25	2567	96.4	501	10	US-09-681-442-8	Sequence 8, Appli
26	2506.5	94.1	476	9	US-09-869-414-6	Sequence 6, Appli
27	2506.5	94.1	476	10	US-09-794-927-6	Sequence 6, Appli
28	2506.5	94.1	476	10	US-09-795-847-6	Sequence 6, Appli
29	2506.5	94.1	476	10	US-09-794-743-6	Sequence 6, Appli
30	2506.5	94.1	476	10	US-09-794-748-6	Sequence 6, Appli
31	2506.5	94.1	476	10	US-09-794-925-6	Sequence 6, Appli
32	2506.5	94.1	476	10	US-09-681-442-6	Sequence 6, Appli
33	2397	90.0	453	9	US-09-869-414-30	Sequence 30, Appl
34	2397	90.0	453	10	US-09-794-927-30	Sequence 30, Appl
35	2397	90.0	453	10	US-09-795-847-30	Sequence 30, Appl
36	2397	90.0	453	10	US-09-794-743-30	Sequence 30, Appl
37	2397	90.0	453	10	US-09-794-748-30	Sequence 30, Appl
38	2397	90.0	453	10	US-09-794-925-30	Sequence 30, Appl
39	2397	90.0	453	10	US-09-681-442-30	Sequence 30, Appl
40	2397	90.0	459	9	US-09-869-414-32	Sequence 32, Appl
41	2397	90.0	459	10	US-09-794-927-32	Sequence 32, Appl
42	2397	90.0	459	10	US-09-795-847-32	Sequence 32, Appl
43	2397	90.0	459	10	US-09-794-743-32	Sequence 32, Appl
44	2397	90.0	459	10	US-09-794-748-32	Sequence 32, Appl
45	2397	90.0	459	10	US-09-794-925-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1

US-09-869-414-4
; Sequence 4, Application US/09869414
; Publication No. US20030077226A1
; GENERAL INFORMATION:
; APPLICANT: Beinkowski et al.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
; FILE REFERENCE: 28341/6280M
; CURRENT APPLICATION NUMBER: US/09/869,414
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-869-414-4

Query Match 100.0%; Score 2664; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAQALPWLMLLMGAGVLPAAHGTQHGI	RLPLRSGLGAPLGLRLPRETDEEPEGRGSF	60
DB	1	MAQALPWLMLLMGAGVLPAAHGTQHGI	RLPLRSGLGAPLGLRLPRETDEEPEGRGSF	60
QY	61	VEVDNLKRGKSCQGYVEMTVGSPQPTLN	ILVDTCSSNFVAGCAAPLPHRYQRLSST	120
DB	61	VEVDNLKRGKSCQGYVEMTVGSPQPTLN	ILVDTCSSNFVAGCAAPLPHRYQRLSST	120
QY	121	YRDLRKGIVVPYTGCKWEGELGTDLVSI	PHGPNVTVRANIAITESDKFFINGSNWEGIL	180

Db 121 YRDLRGVYVYTOCKWEGELGTLVSI PHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDLPEFPFDSLVKQTHVFNLFSLHLCGAGFPPLNOSSEVLASVGGSMIIGGI 240
181 GLAYAEIARPDSDLPEFPFDSLVKQTHVFNLFSLHLCGAGFPPLNOSSEVLASVGGSMIIGGI 240
241 DHSLYTGSLSWTPPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
241 DHSLYTGSLSWTPPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
301 VFEAAVKSIAKASSTKPFDPGFWLCEQLVQVQAGTTPWNIIPVVISLYLMGEVTVNOSFRIT 360
301 VFEAAVKSIAKASSTKPFDPGFWLCEQLVQVQAGTTPWNIIPVVISLYLMGEVTVNOSFRIT 360
361 ILPOOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
361 ILPOOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALEFMLPLCLMVCQW 480
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALEFMLPLCLMVCQW 480
QY 481 RCLRLRQOHHDDFADDISLLK 501
Db 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 2
US-09-794-927-4
; Sequence 4, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
QY 61 VEMVDNLRKSGQGYVYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120
Db 61 VEMVDNLRKSGQGYVYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120

QY 121 YRDLRGVYVYTOCKWEGELGTLVSI PHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
Db 121 YRDLRGVYVYTOCKWEGELGTLVSI PHGPNVTVRANIAAITESSDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPEFPFDSLVKQTHVFNLFSLHLCGAGFPPLNOSSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLPEFPFDSLVKQTHVFNLFSLHLCGAGFPPLNOSSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSLSWTPPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
Db 241 DHSLYTGSLSWTPPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDVSSTTNLRPKK 300
QY 301 VFEAAVKSIAKASSTKPFDPGFWLCEQLVQVQAGTTPWNIIPVVISLYLMGEVTVNOSFRIT 360
Db 301 VFEAAVKSIAKASSTKPFDPGFWLCEQLVQVQAGTTPWNIIPVVISLYLMGEVTVNOSFRIT 360
QY 361 ILPOOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
Db 361 ILPOOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALEFMLPLCLMVCQW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDDESTLMTIAYVMAAICALEFMLPLCLMVCQW 480
QY 481 RCLRLRQOHHDDFADDISLLK 501
Db 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 3
US-09-795-847-4
; Sequence 4, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrichson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
Db 1 MAQALPWLMLLMGAGVLPAGHTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEGRGGSF 60
QY 61 VEMVDNLRKSGQGYVYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQRLSST 120

Db 61 VEMVDNLGRSGGQYVEMTVGSPQTNLILVDTSNFAVGAAPHFLHRYQRLSST 120
121 YRDLRKGVYVPTGCKWEGELGTLVSIPIHGPNTVVRANITAAITSDKFFINGSNWEGIL 180
121 YRDLRKGVYVPTGCKWEGELGTLVSIPIHGPNTVVRANITAAITSDKFFINGSNWEGIL 180
121 YRDLRKGVYVPTGCKWEGELGTLVSIPIHGPNTVVRANITAAITSDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
241 DHSLYTGLWYTPTRREWEYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
241 DHSLYTGLWYTPTRREWEYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
301 VFEAAVKSIAASTERKFPDGFGLGEOLVCMQAGTTPWNIPFPVLSLYLMGEVTVNOSFRIT 360
301 VFEAAVKSIAASTERKFPDGFGLGEOLVCMQAGTTPWNIPFPVLSLYLMGEVTVNOSFRIT 360
361 ILPOOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
361 ILPOOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
481 RCLRCLRQHQHDFADDDISLLK 501
481 RCLRCLRQHQHDFADDDISLLK 501

RESULT 4

US-09-794-743-4
; Sequence 4, Application US/09794743
; Patent No. US2001021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280BC
; CURRENT APPLICATION NUMBER: US/09/794,743
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEPRGRGSF 60
Db 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEPRGRGSF 60

Qy 61 VEMVDNLGRSGGQYVEMTVGSPQTNLILVDTSNFAVGAAPHFLHRYQRLSST 120
Db 61 VEMVDNLGRSGGQYVEMTVGSPQTNLILVDTSNFAVGAAPHFLHRYQRLSST 120
Qy 121 YRDLRKGVYVPTGCKWEGELGTLVSIPIHGPNTVVRANITAAITSDKFFINGSNWEGIL 180
Db 121 YRDLRKGVYVPTGCKWEGELGTLVSIPIHGPNTVVRANITAAITSDKFFINGSNWEGIL 180
Qy 181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDSDLEPFDSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
Qy 241 DHSLYTGLWYTPTRREWEYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Db 241 DHSLYTGLWYTPTRREWEYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIAASTERKFPDGFGLGEOLVCMQAGTTPWNIPFPVLSLYLMGEVTVNOSFRIT 360
Db 301 VFEAAVKSIAASTERKFPDGFGLGEOLVCMQAGTTPWNIPFPVLSLYLMGEVTVNOSFRIT 360
Qy 361 ILPOOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Db 361 ILPOOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFYVVDRAKRRIGFAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
Qy 481 RCLRCLRQHQHDFADDDISLLK 501
Db 481 RCLRCLRQHQHDFADDDISLLK 501

RESULT 5

US-09-794-748-4
; Sequence 4, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQALPWLMMGAGVLPFAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPEEPRGRGSF 60

1 MAQALPWL... 60
1 MAQALPWL... 60
61 VEMVDNL... 120
61 VEMVDNL... 120
121 YRDLRKG... 180
121 YRDLRKG... 180
181 GLAYAEI... 240
181 GLAYAEI... 240
241 DHSLYT... 300
241 DHSLYT... 300
301 VFEAAV... 360
301 VFEAAV... 360
361 ILPQOY... 420
361 ILPQOY... 420
421 HVHDEF... 480
421 HVHDEF... 480
481 RCLRL... 501
481 RCLRL... 501

RESULT 7
US-09-681-442-4
; Sequence 4, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-681-442-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAQALPWL... 60
61 VEMVDNL... 120
61 VEMVDNL... 120
121 YRDLRKG... 180
121 YRDLRKG... 180
181 GLAYAEI... 240
181 GLAYAEI... 240
241 DHSLYT... 300
241 DHSLYT... 300
301 VFEAAV... 360
301 VFEAAV... 360
361 ILPQOY... 420
361 ILPQOY... 420
421 HVHDEF... 480
421 HVHDEF... 480
481 RCLRL... 501
481 RCLRL... 501

RESULT 6
US-09-794-925-4
; Sequence 4, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR
; FILE REFERENCE: 28341/6280HI
; CURRENT FILING DATE: 2001-02-27
; PRIOR FILING DATE: 1999-10-13
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1999-09-23
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-925-4

Query Match 100.0%; Score 2664; DB 10; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.2e-241;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 DB 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 QY 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
 DB 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
 QY 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 DB 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 DB 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 QY 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300
 DB 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300
 QY 301 VFEAAVKSIKAASSTKFPDGFWMGEQLVCMQAGTTPWNIPVVISLYLMGEVTVNSFRIT 360
 DB 301 VFEAAVKSIKAASSTKFPDGFWMGEQLVCMQAGTTPWNIPVVISLYLMGEVTVNSFRIT 360
 QY 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARKRIGFAVSAC 420
 DB 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARKRIGFAVSAC 420
 QY 421 HVHDEFRTAAVGPVETLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOM 480
 DB 421 HVHDEFRTAAVGPVETLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOM 480
 QY 481 RCLRLRQOHHDFADDISLLK 501
 DB 481 RCLRLRQOHHDFADDISLLK 501

RESULT 8

US-10-032-818-4
 ; Sequence 4, Application US/10032818
 ; Publication No. US20030092629A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Koelsch, Gerald
 ; APPLICANT: Ghosh, Arun K.
 ; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
 ; FILE REFERENCE: 2932 1006-007
 ; CURRENT APPLICATION NUMBER: US/10/032,818
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: US 60/275,756
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: US 60/258,705
 ; PRIOR FILING DATE: 2000-12-28
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-10-032-818-4

Query Match 99.7%; Score 2656; DB 9; Length 501;
 Best Local Similarity 99.8%; Pred. No. 2.9e-240;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 DB 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 QY 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
 DB 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120

QY 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 DB 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 DB 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 QY 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300
 DB 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300
 QY 301 VFEAAVKSIKAASSTKFPDGFWMGEQLVCMQAGTTPWNIPVVISLYLMGEVTVNSFRIT 360
 DB 301 VFEAAVKSIKAASSTKFPDGFWMGEQLVCMQAGTTPWNIPVVISLYLMGEVTVNSFRIT 360
 QY 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARKRIGFAVSAC 420
 DB 361 ILPOQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARKRIGFAVSAC 420
 QY 421 HVHDEFRTAAVGPVETLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOM 480
 DB 421 HVHDEFRTAAVGPVETLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOM 480
 QY 481 RCLRLRQOHHDFADDISLLK 501
 DB 481 RCLRLRQOHHDFADDISLLK 501

RESULT 9

US-10-214-932-104
 ; Sequence 104, Application US/10214932
 ; Publication No. US20030100707A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HWANG, Inhwan
 ; APPLICANT: KIM, Dae Heon
 ; APPLICANT: LEE, Yong Jik
 ; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
 ; FILE REFERENCE: APO2/US
 ; CURRENT APPLICATION NUMBER: US/10/214,932
 ; CURRENT FILING DATE: 2002-08-08
 ; NUMBER OF SEQ ID NOS: 133
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 104
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-214-932-104

Query Match 99.7%; Score 2656; DB 9; Length 501;
 Best Local Similarity 99.8%; Pred. No. 2.9e-240;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 DB 1 MAQALPWLMLMAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEGRGSGF 60
 QY 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
 DB 61 VEMVDNLKCKSGQGYVEMTGPOTNLILVDTGSSNFVAGCAAPHPFLHRYQRLSST 120
 QY 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 DB 121 YDLRKGYYVPYTOCKWEGELGTDLSVPHGPNVTVRANIAAITESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 DB 181 GLAYAEIARPDSDLPPFDLSVKQTHVFNPLFSLHLCGAGFPLNQSEVLASVGGSMIIGI 240
 QY 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300
 DB 241 DLSLTGSLWYTPIRREWYEVIIIVRVEINGODLKMDCKEYNDKSIYDVSQTTNLRPKK 300

us-09-869-414a-4.rapb

Fri Jun 6 08:30:59 2003

QY 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
 Db 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
 QY 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 Db 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 QY 481 RCLRLRQOHHDDFADDISLLK 501
 Db 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 11

US-09-795-903A-2
 ; Sequence 2, Application US/09795903A
 ; Patent No. US20020164760A1
 ; GENERAL INFORMATION:

APPLICANT: Tang, Jordan J.N.
 APPLICANT: Koelsch, Gerald
 TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
 FILE REFERENCE: OMRF 179
 CURRENT APPLICATION NUMBER: US/09795.903A
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 09/604,608
 PRIOR FILING DATE: 2000-06-27
 PRIOR APPLICATION NUMBER: 60/168,060
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: 60/177,836
 PRIOR FILING DATE: 2000-01-25
 PRIOR APPLICATION NUMBER: 60/178,368
 PRIOR FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: 60/210,292
 PRIOR FILING DATE: 2000-06-08
 NUMBER OF SEQ ID NOS: 31
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 2
 LENGTH: 488
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Purified Memapsin 2
 OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
 OTHER INFORMATION: residues:
 OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
 OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
 OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
 OTHER INFORMATION: inhibitor
 OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
 OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
 OTHER INFORMATION: 220-224 are N-lobe Beta Strands
 OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
 OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
 OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
 OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
 OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
 OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
 OTHER INFORMATION: and 427-431 are C-lobe Helices

Query Match 96.9%; Score 2582; DB 9; Length 488;

Best Local Similarity 99.8%; Pred. No. 2.4e-233;
 Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPPGRGSGFVEMVMDNLKSGQ 73
 Db 1 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPPGRGSGFVEMVMDNLKSGQ 60
 QY 74 GYVYEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQORQLSSTYRDLRGVVPYT 133
 Db 61 GYVYEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQORQLSSTYRDLRGVVPYT 120

QY 301 VFEAAVKSIAASSTKPDGFWLGEOLVCWQAGTTPWNIPIPVISLYLMGEVTNQSFRT 360
 Db 301 VFEAAVKSIAASSTKPDGFWLGEOLVCWQAGTTPWNIPIPVISLYLMGEVTNQSFRT 360
 QY 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
 Db 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420
 QY 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 Db 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFMLPLCLMVCOM 480
 QY 481 RCLRLRQOHHDDFADDISLLK 501
 Db 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 10

US-09-969-671A-2
 ; Sequence 2, Application US/09969671A
 ; Publication No. US20030036112A1
 ; GENERAL INFORMATION:

APPLICANT: CHAPMAN, CONRAD G.
 APPLICANT: MURPHY, KAY
 APPLICANT: POWELL, DAVID J.
 APPLICANT: SMITH, TRUDI S.
 TITLE OF INVENTION: ASP2
 FILE REFERENCE: GH-70368-D1
 CURRENT APPLICATION NUMBER: US/09/969,671A
 CURRENT FILING DATE: 2001-10-03
 PRIOR APPLICATION NUMBER: UK 9701684.4
 PRIOR FILING DATE: 1997-01-28
 PRIOR APPLICATION NUMBER: 09/009,191
 PRIOR FILING DATE: 1998-01-20
 PRIOR APPLICATION NUMBER: 09/694,200
 PRIOR FILING DATE: 2000-10-23
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 501
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-969-671A-2

Query Match 99.5%; Score 2650; DB 9; Length 501;

Best Local Similarity 99.6%; Pred. No. 1.1e-239;
 Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPPGRGSGF 60
 Db 1 MAQALPWLMLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPPGRGSGF 60
 QY 61 VEMVMDNLKSGQGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQORQLSST 120
 Db 61 VEMVMDNLKSGQGGYVEMTVGSPPTNLILVDTGSSNFVAGAAPHPFLHRYQORQLSST 120
 QY 121 YRDLRGVVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAIIESDKFFINGSNWEGIL 180
 Db 121 YRDLRGVVPYTOCKWEGELGTLVSIPIHGPNTVTRANIAIIESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARPDLSLEPFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 Db 181 GLAYAEIARPDLSLEPFFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
 QY 241 DSHLYTGSLSWYTPIREWYEVIIYRVEINGQDLKMDCKEYNDKSIYDSGTTLNRLPKK 300
 Db 241 DSHLYTGSLSWYTPIREWYEVIIYRVEINGQDLKMDCKEYNDKSIYDSGTTLNRLPKK 300
 QY 301 VFEAAVKSIAASSTKPDGFWLGEOLVCWQAGTTPWNIPIPVISLYLMGEVTNQSFRT 360
 Db 301 VFEAAVKSIAASSTKPDGFWLGEOLVCWQAGTTPWNIPIPVISLYLMGEVTNQSFRT 360

QY 134 QKWEGLGTLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGILGLAYAEIARPDPS 193
 DB 121 QKWEGLGTLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGILGLAYAEIARPDPS 180
 QY 194 LEPPFDSLVKQTHVPNLFSLHLCGAGFPLNOSVLAASVGGSMITGGIDHSLYTGSLWYTP 253
 DB 181 LEPPFDSLVKQTHVPNLFSLHLCGAGFPLNOSVLAASVGGSMITGGIDHSLYTGSLWYTP 240
 QY 254 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKKVFEEAAVKSIRKAS 313
 DB 241 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKKVFEEAAVKSIRKAS 300
 QY 314 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFPVLSLYLMGEVNTNOSFRITILPOOYLRLPVEDV 373
 DB 301 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFPVLSLYLMGEVNTNOSFRITILPOOYLRLPVEDV 360
 QY 374 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 433
 DB 361 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 420
 QY 434 PFVTLDMEDCCYNIPQTDDESTLMTIAYVMAAICAFMLPLCLMVCOMRCLRLRQHQHDDF 493
 DB 421 PFVTLDMEDCCYNIPQTDDESTLMTIAYVMAAICAFMLPLCLMVCOMRCLRLRQHQHDDF 480
 QY 494 ADDISLLK 501
 DB 481 ADDISLLK 488

RESULT 12

US-10-032-818-2
 ; Sequence 2, Application US/10032818
 ; Publication No. US20030092629A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Koelsch, Gerald
 ; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
 ; FILE REFERENCE: 2932.1006-007
 ; CURRENT APPLICATION NUMBER: US/10/032,818
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: US 60/258,705
 ; PRIOR FILING DATE: 2000-12-28
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-032-818-2

Query Match 96.9%; Score 2582; DB 9; Length 488;
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;
 Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSLGGAPLGLRLPRETDEEPEEPRGRGSFVEMVDNLGRKSGQ 73
 DB 1 AGVLPAGHTQHGIRLPLRSLGGAPLGLRLPRETDEEPEEPRGRGSFVEMVDNLGRKSGQ 60
 QY 74 GYVEMTVGSPQTILNLDVGTSSNFVGAAPHFLHRYQRLSSTYRDLRKGYYVYPY 133
 DB 61 GYVEMTVGSPQTILNLDVGTSSNFVGAAPHFLHRYQRLSSTYRDLRKGYYVYPY 120
 QY 134 QKWEGLGTLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGILGLAYAEIARPDPS 193
 DB 121 QKWEGLGTLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGILGLAYAEIARPDPS 180
 QY 194 LEPPFDSLVKQTHVPNLFSLHLCGAGFPLNOSVLAASVGGSMITGGIDHSLYTGSLWYTP 253
 DB 181 LEPPFDSLVKQTHVPNLFSLHLCGAGFPLNOSVLAASVGGSMITGGIDHSLYTGSLWYTP 240

QY 254 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKKVFEEAAVKSIRKAS 313
 DB 241 IRREWEYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTGTTNLRPLPKKVFEEAAVKSIRKAS 300
 QY 314 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFPVLSLYLMGEVNTNOSFRITILPOOYLRLPVEDV 373
 DB 301 STEKPPDGFVLGEOQLVCWQAGTTPWNIPFPVLSLYLMGEVNTNOSFRITILPOOYLRLPVEDV 360
 QY 374 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 433
 DB 361 ATSDODCYKFAISOSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEG 420
 QY 434 PFVTLDMEDCCYNIPQTDDESTLMTIAYVMAAICAFMLPLCLMVCOMRCLRLRQHQHDDF 493
 DB 421 PFVTLDMEDCCYNIPQTDDESTLMTIAYVMAAICAFMLPLCLMVCOMRCLRLRQHQHDDF 480
 QY 494 ADDISLLK 501
 DB 481 ADDISLLK 488

RESULT 13

US-09-796-264-2
 ; Sequence 2, Application US/09796264
 ; Patent No. US20020049303A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Jordan J.N.
 ; APPLICANT: Lin, Xinli
 ; APPLICANT: Koelsch, Gerald
 ; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
 ; FILE REFERENCE: OMRF 179
 ; CURRENT APPLICATION NUMBER: US/09/796,264
 ; CURRENT FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 09/604,608
 ; PRIOR FILING DATE: 2000-06-27
 ; PRIOR APPLICATION NUMBER: 60/168,060
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: 60/177,836
 ; PRIOR FILING DATE: 2000-01-25
 ; PRIOR APPLICATION NUMBER: 60/178,368
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: 60/210,292
 ; PRIOR FILING DATE: 2000-06-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Purified Memapsin 2
 ; OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
 ; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
 ; OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
 ; OTHER INFORMATION: 376-377 are residues in contact with the OM99-2
 ; OTHER INFORMATION: inhibitor
 ; OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
 ; OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
 ; OTHER INFORMATION: 220-224 are N-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices
 ; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
 ; OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
 ; OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
 ; OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
 ; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
 ; OTHER INFORMATION: and 427-431 are C-lobe Helices
 US-09-796-264-2

Query Match 96.9%; Score 2582; DB 10; Length 488;
 Best Local Similarity 99.8%; Pred. No. 2.4e-233;

Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSGFVEMVDNLKRGSGQ 73

DB 1 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSGFVEMVDNLKRGSGQ 60

QY 74 GYVEMTVGSPQTLNLVDTGSSNFAGVGAAPHPFLHRYQRLSSTYRDLRKGVPYPT 133

DB 61 GYVEMTVGSPQTLNLVDTGSSNFAGVGAAPHPFLHRYQRLSSTYRDLRKGVPYPT 120

QY 134 QGKWEGLGTDLSIPHGPNVTVRANITAAITSDKFFINGSNWEGILGLAYAEIARPDOS 193

DB 121 QGKWEGLGTDLSIPHGPNVTVRANITAAITSDKFFINGSNWEGILGLAYAEIARPDOS 180

QY 194 LEPPFDSLKVKTHVNPFLSHLCGAGFPNQSSEVLASVGSMTIGDHSLYTGSWYTP 253

DB 181 LEPPFDSLKVKTHVNPFLSHLCGAGFPNQSSEVLASVGSMTIGDHSLYTGSWYTP 240

QY 254 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTTNLRPKKVFEEAAVKSIAAS 313

DB 241 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTTNLRPKKVFEEAAVKSIAAS 300

QY 314 STEKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVNTNQSFRITILPQOYLRPVEDV 373

DB 301 STEKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVNTNQSFRITILPQOYLRPVEDV 360

QY 374 ATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRKRGFAVSACHVDEFTAAVEG 433

DB 361 ATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRKRGFAVSACHVDEFTAAVEG 420

QY 434 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQWRCRLRQHQHDDF 493

DB 421 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQWRCRLRQHQHDDF 480

QY 494 ADDISLLK 501

DB 481 ADDISLLK 488

RESULT 14

US-09-845-226-2

Sequence 2, Application US/09845226

Patent No. US20020115600A1

GENERAL INFORMATION:

APPLICANT: Tang, Jordan J.N.

APPLICANT: Hong, Lin

APPLICANT: Ghosh, Arun K.

TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof

FILE REFERENCE: OMRP 182

CURRENT APPLICATION NUMBER: US/09/845,226

CURRENT FILING DATE: 2001-04-30

PRIOR APPLICATION NUMBER: 09/603,713

PRIOR FILING DATE: 2000-06-27

PRIOR APPLICATION NUMBER: 60/168,060

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: 60/177,836

PRIOR FILING DATE: 2000-01-25

PRIOR APPLICATION NUMBER: 60/178,368

PRIOR FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: 60/210,292

PRIOR FILING DATE: 2000-06-08

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 488

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Purified Memapsin 2

OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide

OTHER INFORMATION: residues

OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

OTHER INFORMATION: 376-377 are residues in contact with the OM99-2

OTHER INFORMATION: inhibitor

OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111, 113-118, 123-134, 143-154, 165-168, 198-202, and 220-224 are N-lobe Beta Strands

OTHER INFORMATION: Amino acids 184-191 and 210-217 are N-lobe Helices

OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260, 273-275, 282-285, 316-318, 331-336, 342-348, 354-357, 366-370, 372-375, 380-383, 390-395, 404-405, and 418-420 are C-lobe Beta Strands

OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387, and 427-431 are C-lobe Helices

US-09-845-226-2

Query Match 96.9%; Score 2582; DB 10; Length 488;

Best Local Similarity 99.8%; Pred. No. 2.4e-233;

Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSGFVEMVDNLKRGSGQ 73

DB 1 AGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEEPEEPPGRGSGFVEMVDNLKRGSGQ 60

QY 74 GYVEMTVGSPQTLNLVDTGSSNFAGVGAAPHPFLHRYQRLSSTYRDLRKGVPYPT 133

DB 61 GYVEMTVGSPQTLNLVDTGSSNFAGVGAAPHPFLHRYQRLSSTYRDLRKGVPYPT 120

QY 134 QGKWEGLGTDLSIPHGPNVTVRANITAAITSDKFFINGSNWEGILGLAYAEIARPDOS 193

DB 121 QGKWEGLGTDLSIPHGPNVTVRANITAAITSDKFFINGSNWEGILGLAYAEIARPDOS 180

QY 194 LEPPFDSLKVKTHVNPFLSHLCGAGFPNQSSEVLASVGSMTIGDHSLYTGSWYTP 253

DB 181 LEPPFDSLKVKTHVNPFLSHLCGAGFPNQSSEVLASVGSMTIGDHSLYTGSWYTP 240

QY 254 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTTNLRPKKVFEEAAVKSIAAS 313

DB 241 IRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIYDSTTNLRPKKVFEEAAVKSIAAS 300

QY 314 STEKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVNTNQSFRITILPQOYLRPVEDV 373

DB 301 STEKFPDGFGLGEOLVQWAGTTPWNIFFPVISLYLMGEVNTNQSFRITILPQOYLRPVEDV 360

QY 374 ATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRKRGFAVSACHVDEFTAAVEG 433

DB 361 ATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVDRAKRKRGFAVSACHVDEFTAAVEG 420

QY 434 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQWRCRLRQHQHDDF 493

DB 421 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQWRCRLRQHQHDDF 480

QY 494 ADDISLLK 501

DB 481 ADDISLLK 488

RESULT 15

US-09-795-903A-3

Sequence 3, Application US/09795903A

Patent No. US20020164760A1

GENERAL INFORMATION:

APPLICANT: Tang, Jordan J.N.

APPLICANT: Lin, Xinni

APPLICANT: Koelsch, Gerald

TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods of Use Thereof

FILE REFERENCE: OMRP 179

CURRENT APPLICATION NUMBER: US/09/795,903A

CURRENT FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: 09/604,608

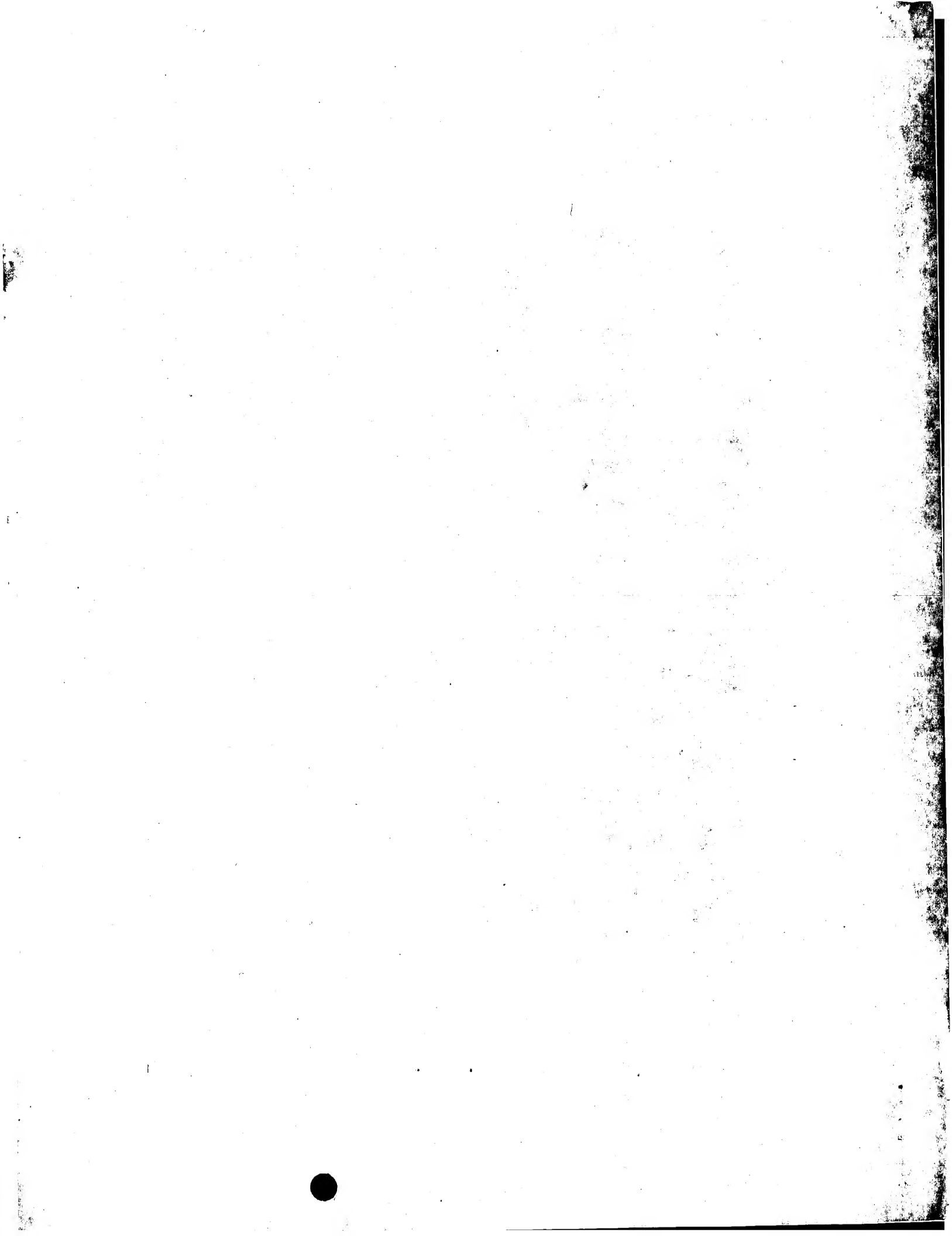
PRIOR FILING DATE: 2000-06-27

PRIOR APPLICATION NUMBER: 60/168,060

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: 60/177,836

; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Pro-memapsin 2
; OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
; OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
; OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
; OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
; OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-795-903A-3
Query Match 96.9%; Score 2582; DB 9; Length 503;
Best Local Similarity 99.8%; Pred. No. 2.5e-233;
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPGRGSGFVEMVDNLRGKSGQ 73
Db 16 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPGRGSGFVEMVDNLRGKSGQ 75
Qy 74 GYVEMTVGSPPTLTNLVDTGSSNFAVGAAPHPFLHRYRQRLSSYRDLRGVYVPYT 133
Db 76 GYVEMTVGSPPTLTNLVDTGSSNFAVGAAPHPFLHRYRQRLSSYRDLRGVYVPYT 135
Qy 134 QGKWEGLGDLVSIPIHGPNTVVRANIAATESDKFFINGSNWEGILGLAYAEIARPDSS 193
Db 136 QGKWEGLGDLVSIPIHGPNTVVRANIAATESDKFFINGSNWEGILGLAYAEIARPDSS 195
Qy 194 LEFFFDLVKQTHVNPNI FSLHLCAGFP LNOSEVLASVGGSMIIGGIDHSLYTGSLWYTP 253
Db 196 LEFFFDLVKQTHVNPNI FSLHLCAGFP LNOSEVLASVGGSMIIGGIDHSLYTGSLWYTP 255
Qy 254 IRREWYEVIIIVVEINGQDLKWDCKEYNYDKSIVDSGTTNLRPLPKKVFEEAAVKSIAAS 313
Db 256 IRREWYEVIIIVVEINGQDLKWDCKEYNYDKSIVDSGTTNLRPLPKKVFEEAAVKSIAAS 315
Qy 314 STEKFPDGFGLGQLVCWQAGTTPWNIFPVISLYLMGEVTNQSFRTILPQQYLRPVEDV 373
Db 316 STEKFPDGFGLGQLVCWQAGTTPWNIFPVISLYLMGEVTNQSFRTILPQQYLRPVEDV 375
Qy 374 ATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAVSACHVHDEFTAAVEG 433
Db 376 ATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKRIGFAVSACHVHDEFTAAVEG 435
Qy 434 PFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQHHDDF 493
Db 436 PFVTLDMEDCGYNIPQTDSTLTMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQHHDDF 495
Qy 494 ADDISLLK 501
Db 496 ADDISLLK 503



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:03:10 ; Search time 38 Seconds
(without alignments)

1756.804 Million cell updates/sec

Title: us-09-869-414a-4

Perfect score: 2664

Sequence: 1 MAQALPWLILMAGVLPAAH.....CLRLRQHQHDFADITSLK 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2664	100.0	501	21	Human aspartyl pro
2	2664	100.0	501	22	Human aspartyl pro
3	2664	100.0	501	22	Human aspartyl pro
4	2664	100.0	501	22	Human aspartyl pro
5	2664	100.0	501	22	Human aspartyl pro
6	2664	100.0	501	22	Human aspartyl pro
7	2664	100.0	501	22	Human aspartyl pro
8	2664	100.0	501	23	Human aspartyl pro
9	2656	99.7	501	21	Human beta-secreta
10	2656	99.7	501	21	Amino acid sequenc

11	2656	99.7	509	23	AA52697	FLAG-tagged human
12	2650	99.5	501	19	AA59807	Amino acid sequenc
13	2588.5	97.2	969	22	ABG09611	Novel human diagno
14	2582	96.9	488	22	AA866572	Human memapsin 2
15	2582	96.9	488	22	AA81334	Memapsin 2 protein
16	2582	96.9	488	23	AAU99488	Human memapsin 2
17	2582	96.9	501	22	AA84948	Mouse aspartic sec
18	2582	96.9	503	22	AA866573	Human pro-memapsin
19	2582	96.9	503	22	AA861335	t7 promoter and ve
20	2582	96.9	503	23	AAU99489	Pro-memapsin 2 enc
21	2569	96.4	501	21	AA94769	Rat beta-secretase
22	2567	96.4	501	21	AA94768	Murine beta-secret
23	2567	96.4	501	21	AA98427	Murine aspartyl pr
24	2567	96.4	501	22	AAE10631	Murine aspartyl pr
25	2567	96.4	501	22	AAE06861	Murine aspartyl pr
26	2567	96.4	501	22	AAU06605	Mouse Aspartyl pro
27	2567	96.4	501	22	AAU07204	Mouse aspartyl pro
28	2567	96.4	501	22	AAE02583	Murine aspartyl pr
29	2567	96.4	501	23	ABB78592	Mouse Asp-2(a) pro
30	2506.5	94.1	476	21	AA98426	Human aspartyl pro
31	2506.5	94.1	476	22	AAE10630	Human aspartyl pro
32	2506.5	94.1	476	22	AAE06860	Human aspartyl pro
33	2506.5	94.1	476	22	AAU06604	Human Aspartyl pro
34	2506.5	94.1	476	22	AAU07203	Human aspartyl pro
35	2506.5	94.1	476	22	AAE02582	Human aspartyl pro
36	2506.5	94.1	476	23	ABB78591	Human Asp-2(b) pro
37	2506.5	94.1	476	23	ABB06410	Human NS protein s
38	2506.5	94.1	476	23	ABB06120	Murine aspartyl pr
39	2420.5	90.9	476	22	AAE06909	Murine aspartyl pr
40	2420.5	90.9	476	22	AAE02619	Active enzyme port
41	2414	90.6	456	21	AA807897	Modified human asp
42	2397	90.0	453	21	AA98438	Human-Asp 2(a) pro
43	2397	90.0	453	22	AAE10642	Human-Asp2(a) delt
44	2397	90.0	453	22	AAE06872	Human-pro-Asp 2(a)
45	2397	90.0	453	22	AAU06616	

ALIGNMENTS

RESULT 1
AA588425
ID AA588425 standard; Protein; 501 AA.
XX
AC AA588425;
XX
DT 03-AUG-2000 (first entry)
XX
DE Human aspartyl protease 2 (a) (Asp2) amino acid sequence.
XX
KW Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 2;
KW Alzheimer's disease; beta secretase site.
XX
OS Homo sapiens.
XX
PN WO200017369-A2.
XX
PD 30-MAR-2000.
XX
PF 23-SEP-1999; 99WO-US20881.
XX
PR 24-SEP-1998; 98US-0101594.
XX
(PHAA) PHARMACIA & UPJOHN CO.
XX
PI Gurney ME, Bienkowski MJ, Heinrichson RL, Parodi LA, Yan R;
XX
DR WPI; 2000-303209/26.
XX
N-PSDB; AAA15662.
XX
PT New enzyme designated human aspartase useful in research into
PT Alzheimer's Disease is capable of cleaving amyloid protein precursor at
PT the beta secretase site to produce amyloid beta peptide

Db 1 MAQALPWLMMGAGVLPAGHTQHGIRLPLRSLGAGPLGLRLPRETDEPEEPPGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDTSNFAVGAAPHPFLHRYYQRLSST 120
Db 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDTSNFAVGAAPHPFLHRYYQRLSST 120
QY 121 YRDLRKGVVYPYTOGKWEGLGTDLVSIPIHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
Db 121 YRDLRKGVVYPYTOGKWEGLGTDLVSIPIHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
QY 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSSTGTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSSTGTTNLRPKK 300
QY 301 VFEAAVKSIKAASSTKFPDGLGELQVLCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Db 301 VFEAAVKSIKAASSTKFPDGLGELQVLCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPQOYLPRVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
Db 361 ILPQOYLPRVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFPLCLMVQW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFPLCLMVQW 480
QY 481 RCLRLRQOHHDFADDISLLK 501
Db 481 RCLRLRQOHHDFADDISLLK 501

RESULT 3

AAE06859
ID AAE06859 standard; Protein; 501 AA.

XX AC AAE06859;

XX DT 23-OCT-2001 (first entry)

XX DE Human aspartyl protease 2a (Hu-Asp2a) protein.

XX KW Human; aspartyl protease 2a; Asp 2a; beta-amyloid precursor protein; App;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nontropic;
KW neuroprotective; antisense therapy; gene therapy;
XX chromosome llq23.3-24.1.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..21

FT Protein /label= signal_peptide

FT Region /note= "Mature human aspartyl protease 2a (Hu-Asp2a)"

FT Domain /note= "Alpha helical spacer region"

FT Domain /label= Transmembrane_domain

FT Domain /label= Cytoplasmic_domain

XX WO200150829-A2.

XX 19-JUL-2001.

XX 09-MAY-2001; 2001WO-IB00799.

XX 09-MAY-2001; 2001WO-IB00799.

XX PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX PI Bienkowski MJ, Gurney ME, Heinrichson RL, Parodi LA, Yan R;
XX DR WPI: 2001-483072/52.
XX DR N-PSDB; AAD13021.
XX DR Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity
XX PS Claim 49; Fig 2; 185pp; English.
XX CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase: Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting
CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is human aspartyl protease 2 (Hu-Asp2), a
CC 'long' form designated as (Hu-Asp2a). Hu-Asp 2 gene is localised on
CC chromosome llq23.3-24.1.
XX SQ Sequence 501 AA;

Query Match 100.0%; Score 2664; DB 22: Length 501;

Best Local Similarity 100.0%; Pred. No. 5.5e-263;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMMGAGVLPAGHTQHGIRLPLRSLGAGPLGLRLPRETDEPEEPPGRGSF 60
Db 1 MAQALPWLMMGAGVLPAGHTQHGIRLPLRSLGAGPLGLRLPRETDEPEEPPGRGSF 60
QY 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDTSNFAVGAAPHPFLHRYYQRLSST 120
Db 61 VEMVDNLRKSGQGYVEMTVGSPPTNLILVDTSNFAVGAAPHPFLHRYYQRLSST 120
QY 121 YRDLRKGVVYPYTOGKWEGLGTDLVSIPIHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
Db 121 YRDLRKGVVYPYTOGKWEGLGTDLVSIPIHGNVTVRANIAAITSDFKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
Db 181 GLAYAEIARPDSDLPPFDLSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240
QY 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSSTGTTNLRPKK 300
Db 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVSSTGTTNLRPKK 300
QY 301 VFEAAVKSIKAASSTKFPDGLGELQVLCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
Db 301 VFEAAVKSIKAASSTKFPDGLGELQVLCWQAGTTPWNIIPVVISLYLMGEVTVNSFRIT 360
QY 361 ILPQOYLPRVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
Db 361 ILPQOYLPRVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGFYVDFDRARKEIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFPLCLMVQW 480

APP or its fragment containing an APP cleavage site recognizable by a mammalian beta-secretase, and further comprising two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian APP or APP fragment. Also included in the invention are methods of identifying modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are useful for treating Alzheimer's disease. APP is useful in methods for identifying inhibitors or modulators of human Asp2 activity and amyloid-beta (Abeta) peptide production. APP is also useful in designing therapeutics for the treatment or prevention of Alzheimer's disease. APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which is associated with increased levels of Abeta processing is useful in assays relating the Alzheimer's research. The expression vector is useful for recombinantly expressing APP. Nucleic acids that hybridize to Asp oligonucleotides are useful as probes or primers. The probes are useful for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and Southern blots. The present sequence is human Asp2(a).

Sequence 501 AA;

Query Match 100.0%; Score 2664; DB 22; Length 501;
Best Local Similarity 100.0%; Pred. No. 5,5e-263;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLWLMGAGVLPAAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPPEPGRGSF 60
DB 1 MAQALPWLWLMGAGVLPAAHGTQHGIRLPLRSGLGGAPLGLRLPRETDEEPPEPGRGSF 60
QY 61 VEMVDNLRGKSGQGYVYVEMTVGSPPTLNILVDVTGSSNFVAVGAAPHPFLHRYQRLSST 120
DB 61 VEMVDNLRGKSGQGYVYVEMTVGSPPTLNILVDVTGSSNFVAVGAAPHPFLHRYQRLSST 120
QY 121 YDLRKGYYVYPTQGWEGELGTDLVSTPHGPNVTVRANTAAITESDKFFINGSNWEGIL 180
DB 121 YDLRKGYYVYPTQGWEGELGTDLVSTPHGPNVTVRANTAAITESDKFFINGSNWEGIL 180
QY 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
DB 181 GLAYAEIARPDSDLPPFDLSLVKQTHVPNLFSLHLCGAGPPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLMDCKEYNDKSIYDVSCTTNLRLPKK 300
DB 241 DHSLYTGSWLYTPIRREWYEVIIIVRVEINGODLMDCKEYNDKSIYDVSCTTNLRLPKK 300
QY 301 VFEAAVKSIAKASSTKEKPDGFWLGEOLVCWQAGTTPWNIPFVLSYLMGEVTVNQSPRIT 360
DB 301 VFEAAVKSIAKASSTKEKPDGFWLGEOLVCWQAGTTPWNIPFVLSYLMGEVTVNQSPRIT 360
QY 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
DB 361 ILPQQYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVDFRARRKIGFAVSAC 420
QY 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFPLCLMLVCQW 480
DB 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFPLCLMLVCQW 480
QY 481 RCLRLRQQHDDFADDSILK 501
DB 481 RCLRLRQQHDDFADDSILK 501

RESULT 5
AAU07202
ID AAU07202 standard; Protein: 501 AA.

XX AC AAU07202;
XX XX 24-OCT-2001 (first entry)
XX DT Human aspartyl protease 2a (Asp-2a).
XX DE Human aspartyl protease 1; Asp-1; neurotrophic; neuroprotective;
XX KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;
XX KW beta-secretase; Alzheimer's disease.

DB 421 HVHDEFTAAVEGPFVTLDMEDCGYNIPTDESTLMTIAYVMAAICALFPLCLMLVCQW 480
QY 481 RCLRLRQQHDDFADDSILK 501
DB 481 RCLRLRQQHDDFADDSILK 501

RESULT 4
AAU06603
ID AAU06603 standard; Protein: 501 AA.

XX AC AAU06603;
XX XX 24-OCT-2001 (first entry)

Human Aspartyl protease 2(a), Asp2(a).

Human; Aspartyl protease; Asp2(a); beta-secretase; neurotrophic; neuroprotective; amyloid protein precursor; APP; Alzheimer's disease; amyloid-beta; Abeta.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..21
FT Peptide /label= Signal peptide
FT Peptide 22..45
FT Peptide /label= Pre_pro-peptide
FT Peptide 46..57
FT Peptide /label= Pro-peptide
FT Protein 57..501
FT Region /label= Mature_Asp2(a)
FT Region 420..454
FT Domain /label= Alpha_helical_spacer_region
FT Domain 455..477
FT Domain /label= Transmembrane_domain
FT Domain 478..501
FT Domain /label= Cytoplasmic_domain

WO200149098-A2.

XX 12-JUL-2001.
XX 09-MAY-2001; 2001WO-IB00798.
XX 09-MAY-2001; 2001WO-IB00798.

(BIEN/) BIENKOWSKI M J.
(GURN/) GURNEY M E.
(HEIN/) HEINRIKSON R L.
(PARO/) PARODI L A.
(YANR/) YAN R.

Bienkowski MJ, Gurney ME, Henrikson RL, Parodi LA, Yan R;
WPI: 2001-502549/55.
N-PSDB; AAS11517.

Novel purified polypeptide comprising fragment of mammalian aspartyl protease 2, lacking Asp2 transmembrane domain and retaining beta secretase activity of Asp2 useful for identifying inhibitors of Asp2 activity

Claim 49; Fig 2; 185pp; English.

The invention relates to a purified polypeptide comprising a fragment of mammalian aspartyl protease (Asp2) protein which lacks the Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide and the fragment retain the beta-secretase activity of the mammalian Asp2 protein. The invention also details polynucleotides for the Asp proteins and vectors expressing them, and a polypeptide (isoform of amyloid protein precursor (APP)) comprising the amino acid sequence of an

FT XX WO200123533-A2. /note= "Peptide #2"

PN XX 05-APR-2001.

PD XX 22-SEP-2000; 2000WO-US26080.

XX XX 23-SEP-1999; 99US-0155493.

PR XX 23-SEP-1999; 99WO-US20881.

PR XX 13-OCT-1999; 99US-0416901.

PR XX 06-DEC-1999; 99US-0169232.

XX XX (PHAA) PHARMACIA & UPJOHN CO.

PA XX Gurney M, Bienkowski MJ;

PI XX WPI; 2001-290516/30.

XX XX N-PSDB; AAD06739.

DR XX

PT XX Enzymes that cleave the alpha-secretase site of the amyloid precursor protein, useful for the treatment of Alzheimer's disease

XX XX Example 2; Fig 2; 189pp; English.

XX XX The present invention relates to enzymes for cleaving the alpha-secretase site of the amyloid precursor protein (APP) and methods of identifying those enzymes. The methods may be used to identify enzymes that may be used to cleave the alpha-secretase cleavage site of the APP protein. The enzymes may be used to treat or modulate the progress of Alzheimer's disease. The present sequence is human aspartyl protease 2a (Asp 2a). Asp 2a has beta-secretase protease activity. Asp 2 gene is located on chromosome 11q23.3-24.1.

XX XX Sequence 501²AA;

XX XX Query Match 100.0%; Score 2664; DB 22; Length 501;

XX XX Best Local Similarity 100.0%; Pred. No. 5.5e-263;

XX XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAQALPWLMLMGAGVLPAGHTGIRLPLRSGLGAPLGLRLPRETDEEPEERGSF 60

Db 1 MAQALPWLMLMGAGVLPAGHTGIRLPLRSGLGAPLGLRLPRETDEEPEERGSF 60

QY 61 VEMVDNLKSGSGGYVEMTVGSPQTLNLVDTGSSNFAYGAAPHLHRYRQLSST 120

61 VEMVDNLKSGSGGYVEMTVGSPQTLNLVDTGSSNFAYGAAPHLHRYRQLSST 120

121 YRDLRKGVYVPTQKWEGLGDLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGIL 180

121 YRDLRKGVYVPTQKWEGLGDLVSIPIHGPNTVVRANTAAITSDKFFINGSNWEGIL 180

QY 181 GLAYAEATRPDLSLEPFDSLVKQTHVPNLFSLHLCAGPPLNQSEVLASVSGSMIIGGI 240

181 GLAYAEATRPDLSLEPFDSLVKQTHVPNLFSLHLCAGPPLNQSEVLASVSGSMIIGGI 240

QY 241 DHSLYTGSLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

241 DHSLYTGSLSWTPIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

QY 301 VFEAAVKSIKAASSTKPPDGFVWLGQLVCHQAGTTPNFPVLSYLMGEVNTNQSFRT 360

301 VFEAAVKSIKAASSTKPPDGFVWLGQLVCHQAGTTPNFPVLSYLMGEVNTNQSFRT 360

QY 361 ILPQOYLARPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKRISFVAVSAC 420

361 ILPQOYLARPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGYVVFDRARRKRISFVAVSAC 420

QY 421 HVHDEFRTAAVEGPPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480

421 HVHDEFRTAAVEGPPVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMWQW 480

QY 481 RCLRLRQOHHDDFADDISLLK 501

Db 481 RCLRLRQOHHDDFADDISLLK 501

RESULT 7

ABB78590 ID ABB78590 standard; Protein; 501 AA.

XX AC ABB78590;

DT 16-JUL-2002 (first entry)

XX XX Human Asp-2(a) protein sequence SEQ ID NO:4.

XX XX Human; Asp-1; Asp-2; aspartyl protease; enzyme; Alzheimer's disease;

KW proteolytic; chromosome 11q23.3-24.1.

OS Homo sapiens.

XX GB2367060-A.

XX 27-MAR-2002.

XX 29-OCT-2001; 2001GB-0025934.

XX 23-SEP-1999; 99US-155493P.

PR 23-SEP-1999; 99US-0404133.

PR 23-SEP-1999; 99WO-US20881.

PR 13-OCT-1999; 99US-0416901.

PR 06-DEC-1999; 99US-169232P.

PR 22-SEP-2000; 2000GB-0023315.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Bienkowski MJ, Gurney M;

XX WPI; 2002-396337/43.

XX N-PSDB; ABL52457.

XX Human aspartyl protease 1 substrates useful in assays to detect

PT aspartyl protease activity, e.g. for the diagnosis of Alzheimer's

XX disease

XX Example 2; Fig 2; 182pp; English.

XX The present invention describes a human aspartyl protease 1 (hu-Asp1)

CC substrate (I) which comprises a peptide of no more than 50 amino acids,

CC and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-

CC Glu-Pro. Also described are: (1) a method (II) for assaying hu-Asp1

CC proteolytic activity, comprising: (a) contacting a hu-Asp1 protein with

CC (I) under acidic conditions; and (b) determining the level of hu-Asp1

CC proteolytic activity; (2) a purified polynucleotide (III) comprising a

CC polynucleotide sequence that hybridises under stringent conditions to the

CC non-coding strand complementary to a defined 1804 nucleotide sequence

CC (see ABL52456) where the nucleotide sequence encodes a polypeptide having

CC Asp1 proteolytic activity and lacks nucleotides encoding a transmembrane

CC domain; (3) a purified polynucleotide (III') comprising a sequence that

CC hybridises under stringent conditions to (III) (the nucleotide sequence

CC encodes a polypeptide further lacking a pro-peptide domain corresponding

CC to amino acids 23-62 of hu-Asp1 (see ABB78589); (4) a vector (IV)

CC comprising (III) or (III') and (5) a host cell (V) transformed or

CC transfected with (III), (III') and/or (IV). The hu-Asp1 protease

CC substrate (I) may be used as an enzyme substrate in assays to detect

CC aspartyl protease activity, (II) and therefore diagnose diseases

CC associated with aberrant hu-Asp1 expression and activity such as

CC Alzheimer's disease. Hu-Asp1 has been localised to chromosome 21, while

CC hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present

CC sequence represents hu-Asp2(a) from the present invention.

XX XX Sequence 501 AA;

XX Query Match 100.0%; Score 2664; DB 23; Length 501;

XX Best Local Similarity 100.0%; Pred. No. 5.5e-263;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MAQALPWLMLWAGVLPAGHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPGRGSF 60	
DB 1 MAQALPWLMLWAGVLPAGHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPGRGSF 60	
QY 61 VEMVNLRLKSGQGYVEMVTGSPPTNLILVDTGSSNFVAGAAHPHFLHRYQRLSST 120	
DB 61 VEMVNLRLKSGQGYVEMVTGSPPTNLILVDTGSSNFVAGAAHPHFLHRYQRLSST 120	
QY 121 YRDLRKGVPVYPTQCKWEGELGTDLVSIIPHGNVTVRANIAAITESDKFFINGSNWEGIL 180	
DB 121 YRDLRKGVPVYPTQCKWEGELGTDLVSIIPHGNVTVRANIAAITESDKFFINGSNWEGIL 180	
QY 181 GLAYAEIARPDSDLPEFFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240	
DB 181 GLAYAEIARPDSDLPEFFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240	
QY 241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVGTTNLRPKK 300	
DB 241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVGTTNLRPKK 300	
QY 301 VFEAAVKSIIKAASSTKFPDGFWLGEOLVCWQAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360	
DB 301 VFEAAVKSIIKAASSTKFPDGFWLGEOLVCWQAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360	
QY 361 ILPQOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420	
DB 361 ILPQOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420	
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVCOW 480	
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVCOW 480	
QY 481 RCLRLRQOHHDDFADDISLLK 501	
DB 481 RCLRLRQOHHDDFADDISLLK 501	
RESULT 8	
ABB06409	
ID ABB06409 standard; Protein; 501 AA.	
XX	
AC ABB06409;	
XX	
DT 31-MAY-2002 (first entry)	
XX	
DE Human aspartyl protease protein sequence SEQ ID NO:2.	
XX	
KW Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP; aspartyl protease; neuroprotective; nontropic; beta-secretase inhibitor; Alzheimer's disease.	
OS Homo sapiens.	
XX	
PN WO200206306-A2.	
XX	
PD 24-JAN-2002.	
XX	
PF 19-JUL-2001; 2001WO-US23035.	
XX	
PR 19-JUL-2000; 2000US-219795P.	
PR 12-MAR-2001; 2001US-275251P.	
XX	
PA (PHAA) PHARMACIA & UPJOHN CO.	
XX	
PI Yan R. Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ; Heinrikson RL;	
XX	
DR WPI; 2002-216995/27.	
XX	
DR N-PSDB; ABL49914.	
XX	
PT Novel substrates for human aspartyl protease useful for identifying	

PT	modulators of beta secretase activity of aspartyl protease for treating Alzheimer's disease
XX	
XX	Claim 63; Page 118-119; 188pp; English.
CC	The present invention describes an isolated peptide (I) comprising a sequence of at least four amino acids, where the peptide is a substrate for conducting aspartyl protease assays. (I) has neuroprotective and nootropic activities, and can be used as an inhibitor of beta-secretase activity. A beta-secretase modulator from the present invention can be used for inhibiting beta-secretase activity in vivo, and in the manufacture of a medicament for the treatment of Alzheimer's disease. Pharmaceutical compositions from the present invention can be used for treating a disease or condition characterised by an abnormal beta-secretase activity. (I) is useful for identifying agents that modulate the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as a core structure to construct derivatives. ABL49914 to ABL49925 and ABB06409 to ABB06593 represent sequences used in the exemplification of the present invention.
XX	
SQ	Sequence 501 AA;
Query Match 100.0%; Score 2664; DB 23; Length 501;	
Best Local Similarity 100.0%; Pred. No. 5.5e-263;	
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MAQALPWLMLWAGVLPAGHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPGRGSF 60	
DB 1 MAQALPWLMLWAGVLPAGHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEPEEPGRGSF 60	
QY 61 VEMVNLRLKSGQGYVEMVTGSPPTNLILVDTGSSNFVAGAAHPHFLHRYQRLSST 120	
DB 61 VEMVNLRLKSGQGYVEMVTGSPPTNLILVDTGSSNFVAGAAHPHFLHRYQRLSST 120	
QY 121 YRDLRKGVPVYPTQCKWEGELGTDLVSIIPHGNVTVRANIAAITESDKFFINGSNWEGIL 180	
DB 121 YRDLRKGVPVYPTQCKWEGELGTDLVSIIPHGNVTVRANIAAITESDKFFINGSNWEGIL 180	
QY 181 GLAYAEIARPDSDLPEFFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240	
DB 181 GLAYAEIARPDSDLPEFFDLSLVKQTHVPLNLSLHLCGAGFPLNQSEVLASVSGSMIIGI 240	
QY 241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVGTTNLRPKK 300	
DB 241 DHSLYTGSLSWYTPIRREWYEVIIIVRVEINGODLKMCKEYNDKSIYDVGTTNLRPKK 300	
QY 301 VFEAAVKSIIKAASSTKFPDGFWLGEOLVCWQAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360	
DB 301 VFEAAVKSIIKAASSTKFPDGFWLGEOLVCWQAGTTPWNIFFVISLYLMGEVNTNOSFRIT 360	
QY 361 ILPQOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420	
DB 361 ILPQOYLPRVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSAC 420	
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVCOW 480	
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPOTDESTLMTIAYVMAAICALFPLCLMVCOW 480	
QY 481 RCLRLRQOHHDDFADDISLLK 501	
DB 481 RCLRLRQOHHDDFADDISLLK 501	
RESULT 9	
AAY94767	
ID AAY94767 standard; Protein; 501 AA.	
XX	
AC AAY94767;	
XX	
DT 12-FEB-2001 (first entry)	
XX	
DE Human beta-secretase amino acid sequence.	
XX	

Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;
Down s syndrome; amyloid angiopathy; gene therapy; neuroprotective.
Homo sapiens.
Key Location/Qualifiers
Peptide 1..45 /label= putative signal peptide
Protein 46..501 /label= Beta-secretase
WO200058479-A1.
05-OCT-2000.
23-MAR-2000; 2000WO-US07755.
26-MAR-1999; 99US-0277229.
(AMGE-) AMGEN INC.
Citron M, Vassar RJ, Bennett BD;
WPI; 2000-594643/56.
N-PSDB; AAA28278.
Isolated beta-secretase nucleic acids and encoded polypeptides, useful
for diagnosis and gene therapy of Alzheimer's disease
Claim 1; Fig 4; 145pp; English.
This invention relates to 3 nucleotide sequences encoding beta-secretase
proteins. Beta-secretase is an enzyme involved in the production of one
of the components of amyloid plaques involved in Alzheimer's disease. The
invention includes an expression vector comprising the nucleotide
sequence, a host cell comprising the expression vector, and a process for
producing the protein through culturing the transformed cells. Also
included in the invention are a polypeptide derivative of the
beta-secretase protein, a fusion protein comprising beta-secretase fused
to a heterologous amino acid sequence, and a method for modulating the
levels of beta-secretase polypeptide in a mammal comprising administering
the polynucleotide sequence. Beta-secretase exhibits neuroprotective and
neurotropic activity. The beta-secretase nucleotide sequence may be used to
map locations of the beta-secretase gene and related genes on chromosomes
and as hybridization probes in diagnostic assays to test for the presence
of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's
syndrome, and amyloid angiopathy. The nucleotide sequence may also be
used as anti-sense inhibitors of beta-secretase expression, in gene
therapy of Alzheimer's disease, and for the identification of compounds
that modulate beta-secretase activity. Antibodies to the beta-secretase
protein may be used for in vitro and in vivo diagnostic purposes to
detect the presence of beta-secretase polypeptide in a body fluid or cell
sample. The present sequence represents the human beta-secretase protein.
Sequence 501 AA;
Query Match 99.7%; Score 2656; DB 21; Length 501;
Best Local Similarity 99.8%; Pred No. 3.6e-262;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 MAQALPWLKMGAGVLPARGTHQIRLPRLRSLGAGAPLGLRLPRETDEEPEEPCRRGSF 60
1 MAQALPWLKMGAGVLPARGTHQIRLPRLRSLGAGAPLGLRLPRETDEEPEEPCRRGSF 60
61 VEMVDNLKSGGGYVEMTVGSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSST*120
61 VEMVDNLKSGGGYVEMTVGSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSST 120
121 YRDLRKGYYVPTQGWEGELGTDLSVIPHPGNVTVVRANAAITSDKFFINGSNWEGIL 180
121 YRDLRKGYYVPTQGWEGELGTDLSVIPHPGNVTVVRANAAITSDKFFINGSNWEGIL 180
181 GLAYAEIARPDSDSLEFFDLSLVKQTHVPLNLSLQLCGAGFPLNQSEVLASVGGSMIIGGI 240

DB 181 GLAYAEIARPDSDSLEFFDLSLVKQTHVPLNLSLQLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNYDKSIYVDSGTTNLRLPKK 300
DB 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNYDKSIYVDSGTTNLRLPKK 300
QY 301 VFEAAVKSITKAASSTKFPDGFGLGOLYCVQWAGITPWNIFPVIISLYLMGEVTVNQSFRIT 360
DB 301 VFEAAVKSITKAASSTKFPDGFGLGOLYCVQWAGITPWNIFPVIISLYLMGEVTVNQSFRIT 360
QY 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVYVDFDRARKRIGFAVSAC 420
DB 361 ILPQOYLRPVEDVATSDQDCYKFAISQSSTGTVMGAVIMEGFYVYVDFDRARKRIGFAVSAC 420
QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAAICALFMLPCLLMVCQW 480
DB 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAAICALFMLPCLLMVCQW 480
QY 481 RCLRLRQOHHDFADDISLLK 501
DB 481 RCLRLRQOHHDFADDISLLK 501
RESULT 10
AAB07896
ID AAB07896 standard; Protein; 501 AA.
XX AAB07896;
AC AAB07896;
XX 14-NOV-2000 (first entry)
DT Amino acid sequence of a human beta-secretase enzyme.
DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;
KW amyloid plaque component; Alzheimer's disease; amyloidogenic disease;
KW inhibitor.
XX Homo sapiens.
XX WO2000047618-A2.
XX 17-AUG-2000.
XX 10-FEB-2000; 2000WO-US03819.
XX 10-FEB-1999; 99US-0119571.
XX 15-JUN-1999; 99US-0139172.
XX (ELAN-) ELAN PHARM INC.
XX Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M;
PI Sinha S, Tatsuno G, Tung J, Wang S, McConlogue L;
XX WPI; 2000-533011/48.
XX N-PSDB; AAA59550, AAA59551.
XX Purified beta-secretase protein used in assays to discover inhibitors
PT which can be used for the treatment of amyloidogenic diseases e.g.
PT Alzheimer's disease
XX Claim 17; Fig 2A; 121pp; English.
XX The specification describes a beta-secretase enzyme. The enzyme cleaves
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This
CC enzyme is therefore implicated in the production of amyloid plaque
CC components which accumulate in the brains of individuals afflicted with
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to
CC a mammalian subject e.g. with Alzheimer's disease or Alzheimer's
CC disease-like pathology to test if they maintain or improve cognitive
CC ability or reduce the plaque burden. The compounds are used for the
CC treatment of amyloidogenic diseases e.g. Alzheimer's disease. The
CC present sequence represents a human beta-secretase enzyme.

361	ILPQOYLRLPVEDVATSQDDCYKFAISQSSTGTVMGAVINMGFYVVDKARKIGFAVSAC	420
421	HVHDEPTAAVEGPFVTLDMEDGYNIPOTDESTLMTIAYVMAATCALFMPLCLMVCOW	480
421		
421	HVHDEPTAAVEGPFVTLDMEDGYNIPOTDESTLMTIAYVMAATCALFMPLCLMVCOW	480
481	RCLRCLRQOHHDFADDISLLK	501
481		
481	RCLRCLRQOHHDFADDISLLK	501
RESULT 13		
ABG09611		
ID	ABG09611 standard; Protein; 969 AA.	
XX	ABG09611;	
XX		
DT	13-FEB-2002 (first entry)	
XX		
DE	Novel human diagnostic protein #9602.	
XX		
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;	
KW	food supplement; medical imaging; diagnostic; genetic disorder.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200175067-A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US08631.	
XX		
PR	31-MAR-2000; 2000US-0540217.	
PR	23-AUG-2000; 2000US-0649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Drmanac RT, Liu C, Tang YT;	
XX		
DR	WPI; 2001-639362/73.	
DR	N-PSDB; AAS73798.	
XX		
XX	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity -	
XX		
PS	Claim 20; SEQ ID NO 39970; 103pp; English.	
XX		
CC	The invention relates to isolated polynucleotide (I) and	
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,	
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome	
CC	and gene mapping, and in recombinant production of (II). The	
CC	polynucleotides are also used in diagnostics as expressed sequence tags	
CC	for identifying expressed genes. (I) is useful in gene therapy techniques	
CC	to restore normal activity of (II) or to treat disease states involving	
CC	(II). (II) is useful for generating antibodies against it, detecting or	
CC	quantitating a polypeptide in tissue, as molecular weight markers and as	
CC	a food supplement. (II) and its binding partners are useful in medical	
CC	imaging of sites expressing (II). (I) and (II) are useful for treating	
CC	disorders involving aberrant protein expression or biological activity.	
CC	The polypeptide and polynucleotide sequences have application in	
CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	and to produce other types of data and products dependent on DNA and	
CC	amino acid sequences. ABG00010-ABG30377 represent novel human	
CC	diagnostic amino acid sequences of the invention.	
CC	Note: The sequence data for this patent did not appear in the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SO	Sequence 969 AA.	

Query Match	97.2%; Score 2588.5; DB 22; Length 969;
Best Local Similarity	98.0%; Pred. No. 8.4e-255;
Matches 492; Conservative 0; Mismatches 9; Indels 1; Gaps 1;	
QY	1 MAQALPWLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60
Db	1 MAQALPWLWAGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEPRGRGSF 60
QY	61 VEMVDNLRGKSGQYVEMVTGSPPTNLILVDGTSSNFVAVGAAPHPFLHRYQRLSST 120
Db	61 VEMVDNLRGKSGQYVEMVTGSPPTNLILVDGTSSNFVAVGAAPHPFLHRYQRLSST 120
QY	121 YRDLKRGVYVPTQGWGEGELGTLVSIPIHGNVTVRANIAAITSDESKFFINGSNWEGIL 180
Db	121 YRDLKRGVYVPTQGWGEGELGTLVSIPIHGNVTVRANIAAITSDESKFFINGSNWEGIL 180
QY	181 GLAYAEIARPDSDLPEFDSLVKQTHVPLNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
Db	181 GLAYAEIARPDSDLPEFDSLVKQTHVPLNLSHLGAGFPLNOSSEVLASVGGSMIIGGI 240
QY	241 DHSLYTGLSWYTPIRREYYEVIIVRVEINGODLKMCKEYNYDKSIVDSGTTNLRPLPK 300
Db	241 DHSLYTGLSWYTPIRREYYEVIIVRVEINGODLKMCKEYNYDKSIVDSGTTNLRPLPK 300
QY	301 VFEAAVKSIKAASSTKEKPDGFWLGEQLVCWQAGTTPWNIFFVISLYLMGEVTVNOSFRIT 360
Db	301 VFEAAVKSIKAASSTKEKPDGFWLGEQLVCWQAGTTPWNIFFVISLYLMGEVTVNOSFRIT 360
QY	361 ILPOQYLRPEVDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVDFRARKRIGFAYSAC 420
Db	361 ILPOQYLRPEVDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVDFRARKRIGFAYSAC 420
QY	421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAAICALFMLPLCLMWQW 480
Db	421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPTQDESTLMTIAYVMAAICALFMLPLCLMWQW 480
QY	481 RCLRLRQOHD-DFADDISLLK 501
Db	481 RCLRLRQOHD-DFADDISLLK 501
RESULT 14	
ID	AAB66572
AC	AAB66572 standard; Protein; 488 AA.
AC	AAB66572;
DT	12-APR-2001 (first entry)
XX	Human memapsin 2.
XX	Human; memapsin 2; neurotrophic; amyloid precursor protein;
XX	APP; memapsin 2 inhibitor; Alzheimer's disease.
OS	Homo sapiens.
XX	WO200100665-A2.
XX	04-JAN-2001.
XX	27-JUN-2000; 2000WO-US17742.
XX	28-JUN-1999; 99US-0141363.
PR	30-NOV-1999; 99US-0168060.
PR	25-JAN-2000; 2000US-0177836.
PR	27-JAN-2000; 2000US-0178368.
PR	08-JUN-2000; 2000US-0210292.
XX	(OKLA-) OKLAHOMA MEDICAL RES FOUND.
PA	(UNII) UNIV ILLINOIS FOUND.
PI	Tang JJN, Hong L, Ghosh AK;

QY 494 ADDISLLK 501
Db 481 ADDISLLK 488

Search completed: June 5, 2003, 16:12:15
Job time : 40 secs

KW Memapsin 2; catalyst; Alzheimer's.
XX OS Homo sapiens.
XX PN WO200100663-A2.
XX PD 04-JAN-2001.
XX PF 27-JUN-2000; 2000WO-US17661.
XX PR 28-JUN-1999; 99US-0141363.
XX PR 30-NOV-1999; 99US-0168060.
XX PR 25-JAN-2000; 2000US-0177836.
XX PR 27-JAN-2000; 2000US-0178368.
XX PR 08-JUN-2000; 2000US-0210292.
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX Tang JUN, Lin X, Koelsch G;
WPI; 2001-102885/11.
XX Purified recombinant catalytically active memapsin 2, used to screen
PT inhibitors of it, which are used to treat and prevent Alzheimer's
PT disease -
XX PT
XX PS
XX Claim 2; Page 73-75; 86pp; English.
XX The present invention relates to a purified recombinant
CC catalytically active memapsin 2. The invention may be used for
CC isolating inhibitors which are used to treat or prevent
CC Alzheimer's disease. The invention may also be used to screen
CC for individuals more genetically prone to develop Alzheimer's
CC disease.
XX
XX Sequence 488 AA;
Query Match 96.9%; Score 2582; DB 22; Length 488;
Best Local Similarity 99.8%; Pred. No. 1.3e-254;
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 14 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSRFVEMVDNLRKSGQ 73
Db 1 AGVLPAGHTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSRFVEMVDNLRKSGQ 60
QY 74 GYVEMTVGSPPTLNLVDYTGSSNFAGVGAAPHLHRYIQRLSTYRDLRKGVVVPYT 133
61 GYVEMTVGSPPTLNLVDYTGSSNFAGVGAAPHLHRYIQRLSTYRDLRKGVVVPYT 120
QY 134 QKWEGELGTDLSIPHPNVTVRANTAAITSDKFFINGSNWEGILGLAYAEIARPPDS 193
Db 121 QKWEGELGTDLSIPHPNVTVRANTAAITSDKFFINGSNWEGILGLAYAEIARPPDS 180
QY 194 LEPPFDSLVTQTHVPNLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSWYTP 253
Db 181 LEPPFDSLVTQTHVPNLFSLHLCAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSWYTP 240
QY 254 IRREMYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKKVFEEAAVKSIAAS 313
Db 241 IRREMYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPKKVFEEAAVKSIAAS 300
QY 314 STEKFPDGFNLGDLVLCWQAGTTPNIFPVLSLYLMGEVTNQSFRTILPQOYLRPVEDV 373
Db 301 STEKFPDGFNLGDLVLCWQAGTTPNIFPVLSLYLMGEVTNQSFRTILPQOYLRPVEDV 360
QY 374 ATSDDCYKFAISQSSGTGMGAVIMEGFYVDFDRKRIGFAVSACHVHDEFRTAAVEG 433
Db 361 ATSDDCYKFAISQSSGTGMGAVIMEGFYVDFDRKRIGFAVSACHVHDEFRTAAVEG 420
QY 434 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQCWRCLRCLRQHQHDF 493
Db 421 PFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICAFMLPLCLMVQCWRCLRCLRQHQHDF 480

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:03:45 ; Search time 22 Seconds
(without alignments)
944.529 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLLMGAGVLPAL.....CLRLCROHDDFADDISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2656	93.7	501	1	BACE_HUMAN
2	2569	96.4	501	1	BACE_RAT
3	2567	96.4	501	1	BACE_MOUSE
4	1173.5	44.1	518	1	BAE2_HUMAN
5	330	12.4	324	1	PEP1_GADMO
6	314.5	11.8	390	1	CATD_BOVIN
7	309	11.6	387	1	PEP1_RABIT
8	307.5	11.5	388	1	PEP4_MACFU
9	305	11.4	367	1	PEP4_CHICK
10	301.5	11.3	383	1	PEPE_CHICK
11	301.5	11.3	396	1	CATE_HUMAN
12	300.5	11.3	412	1	CATD_HUMAN
13	299	11.2	387	1	PEP2_RABIT
14	298	11.2	387	1	PEP4_RABIT
15	297	11.1	407	1	CATD_RAT
16	295	11.1	391	1	CATE_CAVPO
17	294.5	11.1	388	1	PEP2_MACFU
18	289	10.8	387	1	PEP3_RABIT
19	288.5	10.8	388	1	PEP4_HUMAN
20	288.5	10.8	388	1	PEP4_MACU
21	288.5	10.8	398	1	CATE_RAT
22	287	10.8	410	1	CATD_MOUSE
23	286.5	10.8	388	1	PEP1_MACFU
24	286	10.7	398	1	CATD_CHICK
25	284.5	10.7	381	1	CHYM_SHEEP
26	281.5	10.6	386	1	PEP4_PIG
27	281	10.5	387	1	PEP4_CALJA
28	280.5	10.5	397	1	CATE_MOUSE
29	276.5	10.4	381	1	CHYM_BOVIN
30	276.5	10.4	396	1	CATE_RABIT
31	274.5	10.3	419	1	CARV_CANAL
32	273.5	10.3	376	1	PAG2_BOVIN
33	273.5	10.3	377	1	PEP2_MACFU

34	273	10.2	388	1	PEPF_RABIT
35	270.5	10.2	381	1	CHYM_CALJA
36	268	10.1	396	1	CARP_NEUCR
37	267	10.0	365	1	CATD_SHEEP
38	266.5	10.0	388	1	PEPC_CALJA
39	266	10.0	394	1	PEPC_CAVPO
40	266	10.0	405	1	CARP_YEAST
41	264.5	9.9	388	1	PEPC_HUMAN
42	264	9.9	388	1	PAG_HORSE
43	262	9.8	496	1	ASPR_ORYSA
44	261.5	9.8	387	1	ASPP_AEDAE
45	261	9.8	345	1	CATD_PIG

ALIGNMENTS

RESULT 1
BACE_HUMAN STANDARD: PRT; 501 AA.
AC P56817; Q9UJT5; Q9BYC1; Q9BYB9;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
protease 2) (ASP 2) (ASP2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2)
GN BACE OR BACE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Brain;
RX MEDLINE=20002972; PubMed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
Treanor J., Rogers G., Citron M.;
RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
the transmembrane aspartic protease BACE.";
RL Science 286:735-741(1999).

[2]
RP SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND
RC CHARACTERIZATION.
RX MEDLINE=20051717; PubMed=10591214;
RA Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavellio R.,
Davis D., Doan M., Dovey H.F., Eriqon N., Hong J., Jacobson-Croak K.,
Jewett N., Kelm P., Knops J., Lieberburg I., Power M., Tan H., Wang S.,
Tatsuno G., Tung J., Schenk D., Seubert P., Suomensaar S.M., Wang S.,
Walker D., Zhao J., McConlogue L., Varghese J.;
RT "Purification and cloning of amyloid precursor protein beta-secretase
from human brain.";
RL Nature 402:537-540(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=20051707; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
Brashier J.R., Strauman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-
secretase activity.";
RL Nature 402:533-537(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=20120043; PubMed=10656250;
RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;

"Identification of a novel aspartic proteinase (Asp 2) as beta-secreta..."
RL Mol-Cell. Neurosci. 14:419-427(1999).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Brain, and Pancreas;
RA Michel B., De Pietri Tonelli D., Zaccchetti D., Keller P.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from
human brain and pancreas";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM C).
RC TISSUE=Pancreas;
RA Zaccchetti D., De Pietri Tonelli D., Schnurbus R.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
human pancreas";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS B; C AND D).
RC TISSUE=Brain;
RL MEDLINE=21408467; PubMed=11516562;
Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site
amyloid precursor protein cleaving enzyme (BACE) and their effect on
amyloid beta-peptide production";
RL Neurosci. Lett. 307:9-12(2001).
RN [8]
RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
RX MEDLINE=20144060; PubMed=10677483;
RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
beta-amyloid precursor protein";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=21950860; PubMed=11953458;
RA Fischer F., Molinari M., Bodendorf U., Paganetti P.;
RT "The disulphide bonds in the catalytic domain of BACE are critical but
not essential for amyloid precursor protein processing activity";
RL J. Neurochem. 80:1079-1088(2002).
CC -1- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms: A/BACE-1A/BACE-501 (shown here),
B/BACE-1B/BACE-1-476; C/BACE-1C/BACE-1-457 and D/BACE-1D/BACE-1-
432; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; AF190725; AAF04142.1;
DR EMBL; AF201468; AAF18982.1;
DR EMBL; AF200343; AAF17079.1;
DR EMBL; AF204943; AAF26367.1;
DR EMBL; AF338816; AAK38374.1;
DR EMBL; AF338817; AAK38375.1;
DR EMBL; AF338818; AAK38376.1;
DR EMBL; AF338819; AAK38377.1;
DR EMBL; AF338820; AAK38378.1;
DR EMBL; AF338821; AAK38379.1;
DR EMBL; AF338822; AAK38380.1;
DR EMBL; AF338823; AAK38381.1;
DR EMBL; AF338824; AAK38382.1;
DR EMBL; AF338825; AAK38383.1;
DR EMBL; AF338826; AAK38384.1;
DR EMBL; AF338827; AAK38385.1;
DR EMBL; AF338828; AAK38386.1;
DR EMBL; AF338829; AAK38387.1;
DR EMBL; AF338830; AAK38388.1;
DR EMBL; AF338831; AAK38389.1;
DR EMBL; AF338832; AAK38390.1;
DR EMBL; AF338833; AAK38391.1;
DR EMBL; AF338834; AAK38392.1;
DR EMBL; AF338835; AAK38393.1;
DR EMBL; AF338836; AAK38394.1;
DR EMBL; AF338837; AAK38395.1;
DR EMBL; AF338838; AAK38396.1;
DR EMBL; AF338839; AAK38397.1;
DR EMBL; AF338840; AAK38398.1;
DR EMBL; AF338841; AAK38399.1;
DR EMBL; AF338842; AAK38400.1;
DR EMBL; AF338843; AAK38401.1;
DR EMBL; AF338844; AAK38402.1;
DR EMBL; AF338845; AAK38403.1;
DR EMBL; AF338846; AAK38404.1;
DR EMBL; AF338847; AAK38405.1;
DR EMBL; AF338848; AAK38406.1;
DR EMBL; AF338849; AAK38407.1;
DR EMBL; AF338850; AAK38408.1;
DR EMBL; AF338851; AAK38409.1;
DR EMBL; AF338852; AAK38410.1;
DR EMBL; AF338853; AAK38411.1;
DR EMBL; AF338854; AAK38412.1;
DR EMBL; AF338855; AAK38413.1;
DR EMBL; AF338856; AAK38414.1;
DR EMBL; AF338857; AAK38415.1;
DR EMBL; AF338858; AAK38416.1;
DR EMBL; AF338859; AAK38417.1;
DR EMBL; AF338860; AAK38418.1;
DR EMBL; AF338861; AAK38419.1;
DR EMBL; AF338862; AAK38420.1;
DR EMBL; AF338863; AAK38421.1;
DR EMBL; AF338864; AAK38422.1;
DR EMBL; AF338865; AAK38423.1;
DR EMBL; AF338866; AAK38424.1;
DR EMBL; AF338867; AAK38425.1;
DR EMBL; AF338868; AAK38426.1;
DR EMBL; AF338869; AAK38427.1;
DR EMBL; AF338870; AAK38428.1;
DR EMBL; AF338871; AAK38429.1;
DR EMBL; AF338872; AAK38430.1;
DR EMBL; AF338873; AAK38431.1;
DR EMBL; AF338874; AAK38432.1;
DR EMBL; AF338875; AAK38433.1;
DR EMBL; AF338876; AAK38434.1;
DR EMBL; AF338877; AAK38435.1;
DR EMBL; AF338878; AAK38436.1;
DR EMBL; AF338879; AAK38437.1;
DR EMBL; AF338880; AAK38438.1;
DR EMBL; AF338881; AAK38439.1;
DR EMBL; AF338882; AAK38440.1;
DR EMBL; AF338883; AAK38441.1;
DR EMBL; AF338884; AAK38442.1;
DR EMBL; AF338885; AAK38443.1;
DR EMBL; AF338886; AAK38444.1;
DR EMBL; AF338887; AAK38445.1;
DR EMBL; AF338888; AAK38446.1;
DR EMBL; AF338889; AAK38447.1;
DR EMBL; AF338890; AAK38448.1;
DR EMBL; AF338891; AAK38449.1;
DR EMBL; AF338892; AAK38450.1;
DR EMBL; AF338893; AAK38451.1;
DR EMBL; AF338894; AAK38452.1;
DR EMBL; AF338895; AAK38453.1;
DR EMBL; AF338896; AAK38454.1;
DR EMBL; AF338897; AAK38455.1;
DR EMBL; AF338898; AAK38456.1;
DR EMBL; AF338899; AAK38457.1;
DR EMBL; AF338900; AAK38458.1;
DR EMBL; AF338901; AAK38459.1;
DR EMBL; AF338902; AAK38460.1;
DR EMBL; AF338903; AAK38461.1;
DR EMBL; AF338904; AAK38462.1;
DR EMBL; AF338905; AAK38463.1;
DR EMBL; AF338906; AAK38464.1;
DR EMBL; AF338907; AAK38465.1;
DR EMBL; AF338908; AAK38466.1;
DR EMBL; AF338909; AAK38467.1;
DR EMBL; AF338910; AAK38468.1;
DR EMBL; AF338911; AAK38469.1;
DR EMBL; AF338912; AAK38470.1;
DR EMBL; AF338913; AAK38471.1;
DR EMBL; AF338914; AAK38472.1;
DR EMBL; AF338915; AAK38473.1;
DR EMBL; AF338916; AAK38474.1;
DR EMBL; AF338917; AAK38475.1;
DR EMBL; AF338918; AAK38476.1;
DR EMBL; AF338919; AAK38477.1;
DR EMBL; AF338920; AAK38478.1;
DR EMBL; AF338921; AAK38479.1;
DR EMBL; AF338922; AAK38480.1;
DR EMBL; AF338923; AAK38481.1;
DR EMBL; AF338924; AAK38482.1;
DR EMBL; AF338925; AAK38483.1;
DR EMBL; AF338926; AAK38484.1;
DR EMBL; AF338927; AAK38485.1;
DR EMBL; AF338928; AAK38486.1;
DR EMBL; AF338929; AAK38487.1;
DR EMBL; AF338930; AAK38488.1;
DR EMBL; AF338931; AAK38489.1;
DR EMBL; AF338932; AAK38490.1;
DR EMBL; AF338933; AAK38491.1;
DR EMBL; AF338934; AAK38492.1;
DR EMBL; AF338935; AAK38493.1;
DR EMBL; AF338936; AAK38494.1;
DR EMBL; AF338937; AAK38495.1;
DR EMBL; AF338938; AAK38496.1;
DR EMBL; AF338939; AAK38497.1;
DR EMBL; AF338940; AAK38498.1;
DR EMBL; AF338941; AAK38499.1;
DR EMBL; AF338942; AAK38500.1;
DR EMBL; AF338943; AAK38501.1;
DR EMBL; AF338944; AAK38502.1;
DR EMBL; AF338945; AAK38503.1;
DR EMBL; AF338946; AAK38504.1;
DR EMBL; AF338947; AAK38505.1;
DR EMBL; AF338948; AAK38506.1;
DR EMBL; AF338949; AAK38507.1;
DR EMBL; AF338950; AAK38508.1;
DR EMBL; AF338951; AAK38509.1;
DR EMBL; AF338952; AAK38510.1;
DR EMBL; AF338953; AAK38511.1;
DR EMBL; AF338954; AAK38512.1;
DR EMBL; AF338955; AAK38513.1;
DR EMBL; AF338956; AAK38514.1;
DR EMBL; AF338957; AAK38515.1;
DR EMBL; AF338958; AAK38516.1;
DR EMBL; AF338959; AAK38517.1;
DR EMBL; AF338960; AAK38518.1;
DR EMBL; AF338961; AAK38519.1;
DR EMBL; AF338962; AAK38520.1;
DR EMBL; AF338963; AAK38521.1;
DR EMBL; AF338964; AAK38522.1;
DR EMBL; AF338965; AAK38523.1;
DR EMBL; AF338966; AAK38524.1;
DR EMBL; AF338967; AAK38525.1;
DR EMBL; AF338968; AAK38526.1;
DR EMBL; AF338969; AAK38527.1;
DR EMBL; AF338970; AAK38528.1;
DR EMBL; AF338971; AAK38529.1;
DR EMBL; AF338972; AAK38530.1;
DR EMBL; AF338973; AAK38531.1;
DR EMBL; AF338974; AAK38532.1;
DR EMBL; AF338975; AAK38533.1;
DR EMBL; AF338976; AAK38534.1;
DR EMBL; AF338977; AAK38535.1;
DR EMBL; AF338978; AAK38536.1;
DR EMBL; AF338979; AAK38537.1;
DR EMBL; AF338980; AAK38538.1;
DR EMBL; AF338981; AAK38539.1;
DR EMBL; AF338982; AAK38540.1;
DR EMBL; AF338983; AAK38541.1;
DR EMBL; AF338984; AAK38542.1;
DR EMBL; AF338985; AAK38543.1;
DR EMBL; AF338986; AAK38544.1;
DR EMBL; AF338987; AAK38545.1;
DR EMBL; AF338988; AAK38546.1;
DR EMBL; AF338989; AAK38547.1;
DR EMBL; AF338990; AAK38548.1;
DR EMBL; AF338991; AAK38549.1;
DR EMBL; AF338992; AAK38550.1;
DR EMBL; AF338993; AAK38551.1;
DR EMBL; AF338994; AAK38552.1;
DR EMBL; AF338995; AAK38553.1;
DR EMBL; AF338996; AAK38554.1;
DR EMBL; AF338997; AAK38555.1;
DR EMBL; AF338998; AAK38556.1;
DR EMBL; AF338999; AAK38557.1;
DR EMBL; AF339000; AAK38558.1;
DR EMBL; AF339001; AAK38559.1;
DR EMBL; AF339002; AAK38560.1;
DR EMBL; AF339003; AAK38561.1;
DR EMBL; AF339004; AAK38562.1;
DR EMBL; AF339005; AAK38563.1;
DR EMBL; AF339006; AAK38564.1;
DR EMBL; AF339007; AAK38565.1;
DR EMBL; AF339008; AAK38566.1;
DR EMBL; AF339009; AAK38567.1;
DR EMBL; AF339010; AAK38568.1;
DR EMBL; AF339011; AAK38569.1;
DR EMBL; AF339012; AAK38570.1;
DR EMBL; AF339013; AAK38571.1;
DR EMBL; AF339014; AAK38572.1;
DR EMBL; AF339015; AAK38573.1;
DR EMBL; AF339016; AAK38574.1;
DR EMBL; AF339017; AAK38575.1;
DR EMBL; AF339018; AAK38576.1;
DR EMBL; AF339019; AAK38577.1;
DR EMBL; AF339020; AAK38578.1;
DR EMBL; AF339021; AAK38579.1;
DR EMBL; AF339022; AAK38580.1;
DR EMBL; AF339023; AAK38581.1;
DR EMBL; AF339024; AAK38582.1;
DR EMBL; AF339025; AAK38583.1;
DR EMBL; AF339026; AAK38584.1;
DR EMBL; AF339027; AAK38585.1;
DR EMBL; AF339028; AAK38586.1;
DR EMBL; AF339029; AAK38587.1;
DR EMBL; AF339030; AAK38588.1;
DR EMBL; AF339031; AAK38589.1;
DR EMBL; AF339032; AAK38590.1;
DR EMBL; AF339033; AAK38591.1;
DR EMBL; AF339034; AAK38592.1;
DR EMBL; AF339035; AAK38593.1;
DR EMBL; AF339036; AAK38594.1;
DR EMBL; AF339037; AAK38595.1;
DR EMBL; AF339038; AAK38596.1;
DR EMBL; AF339039; AAK38597.1;
DR EMBL; AF339040; AAK38598.1;
DR EMBL; AF339041; AAK38599.1;
DR EMBL; AF339042; AAK38600.1;
DR EMBL; AF339043; AAK38601.1;
DR EMBL; AF339044; AAK38602.1;
DR EMBL; AF339045; AAK38603.1;
DR EMBL; AF339046; AAK38604.1;
DR EMBL; AF339047; AAK38605.1;
DR EMBL; AF339048; AAK38606.1;
DR EMBL; AF339049; AAK38607.1;
DR EMBL; AF339050; AAK38608.1;
DR EMBL; AF339051; AAK38609.1;
DR EMBL; AF339052; AAK38610.1;
DR EMBL; AF339053; AAK38611.1;
DR EMBL; AF339054; AAK38612.1;
DR EMBL; AF339055; AAK38613.1;
DR EMBL; AF339056; AAK38614.1;
DR EMBL; AF339057; AAK38615.1;
DR EMBL; AF339058; AAK38616.1;
DR EMBL; AF339059; AAK38617.1;
DR EMBL; AF339060; AAK38618.1;
DR EMBL; AF339061; AAK38619.1;
DR EMBL; AF339062; AAK38620.1;
DR EMBL; AF339063; AAK38621.1;
DR EMBL; AF339064; AAK38622.1;
DR EMBL; AF339065; AAK38623.1;
DR EMBL; AF339066; AAK38624.1;
DR EMBL; AF339067; AAK38625.1;
DR EMBL; AF339068; AAK38626.1;
DR EMBL; AF339069; AAK38627.1;
DR EMBL; AF339070; AAK38628.1;
DR EMBL; AF339071; AAK38629.1;
DR EMBL; AF339072; AAK38630.1;
DR EMBL; AF339073; AAK38631.1;
DR EMBL; AF339074; AAK38632.1;
DR EMBL; AF339075; AAK38633.1;
DR EMBL; AF339076; AAK38634.1;
DR EMBL; AF339077; AAK38635.1;
DR EMBL; AF339078; AAK38636.1;
DR EMBL; AF339079; AAK38637.1;
DR EMBL; AF339080; AAK38638.1;
DR EMBL; AF339081; AAK38639.1;
DR EMBL; AF339082; AAK38640.1;
DR EMBL; AF339083; AAK38641.1;
DR EMBL; AF339084; AAK38642.1;
DR EMBL; AF339085; AAK38643.1;
DR EMBL; AF339086; AAK38644.1;
DR EMBL; AF339087; AAK38645.1;
DR EMBL; AF339088; AAK38646.1;
DR EMBL; AF339089; AAK38647.1;
DR EMBL; AF339090; AAK38648.1;
DR EMBL; AF339091; AAK38649.1;
DR EMBL; AF339092; AAK38650.1;
DR EMBL; AF339093; AAK38651.1;
DR EMBL; AF339094; AAK38652.1;
DR EMBL; AF339095; AAK38653.1;
DR EMBL; AF339096; AAK38654.1;
DR EMBL; AF339097; AAK38655.1;
DR EMBL; AF339098; AAK38656.1;
DR EMBL; AF339099; AAK38657.1;
DR EMBL; AF339100; AAK38658.1;
DR EMBL; AF339101; AAK38659.1;
DR EMBL; AF339102; AAK38660.1;
DR EMBL; AF339103; AAK38661.1;
DR EMBL; AF339104; AAK38662.1;
DR EMBL; AF339105; AAK38663.1;
DR EMBL; AF339106; AAK38664.1;
DR EMBL; AF339107; AAK38665.1;
DR EMBL; AF339108; AAK38666.1;
DR EMBL; AF339109; AAK38667.1;
DR EMBL; AF339110; AAK38668.1;
DR EMBL; AF339111; AAK38669.1;
DR EMBL; AF339112; AAK38670.1;
DR EMBL; AF339113; AAK38671.1;
DR EMBL; AF339114; AAK38672.1;
DR EMBL; AF339115; AAK38673.1;
DR EMBL; AF339116; AAK38674.1;
DR EMBL; AF339117; AAK38675.1;
DR EMBL; AF339118; AAK38676.1;
DR EMBL; AF339119; AAK38677.1;
DR EMBL; AF339120; AAK38678.1;
DR EMBL; AF339121; AAK38679.1;
DR EMBL; AF339122; AAK38680.1;
DR EMBL; AF339123; AAK38681.1;
DR EMBL; AF339124; AAK38682.1;
DR EMBL; AF339125; AAK38683.1;
DR EMBL; AF339126; AAK38684.1;
DR EMBL; AF339127; AAK38685.1;
DR EMBL; AF339128; AAK38686.1;
DR EMBL; AF339129; AAK38687.1;
DR EMBL; AF339130; AAK38688.1;
DR EMBL; AF339131; AAK38689.1;
DR EMBL; AF339132; AAK38690.1;
DR EMBL; AF339133; AAK38691.1;
DR EMBL; AF339134; AAK38692.1;
DR EMBL; AF339135; AAK38693.1;
DR EMBL; AF339136; AAK38694.1;
DR EMBL; AF339137; AAK38695.1;
DR EMBL; AF339138; AAK38696.1;
DR EMBL; AF339139; AAK38697.1;
DR EMBL; AF339140; AAK38698.1;
DR EMBL; AF339141; AAK38699.1;
DR EMBL; AF339142; AAK38700.1;
DR EMBL; AF339143; AAK38701.1;
DR EMBL; AF339144; AAK38702.1;
DR EMBL; AF339145; AAK38703.1;
DR EMBL; AF339146; AAK38704.1;
DR EMBL; AF339147; AAK38705.1;
DR EMBL; AF339148; AAK38706.1;
DR EMBL; AF339149; AAK38707.1;
DR EMBL; AF339150; AAK38708.1;
DR EMBL; AF339151; AAK38709.1;
DR EMBL; AF339152; AAK38710.1;
DR EMBL; AF339153; AAK38711.1;
DR EMBL; AF339154; AAK38712.1;
DR EMBL; AF339155; AAK38713.1;
DR EMBL; AF339156; AAK38714.1;
DR EMBL; AF339157; AAK38715.1;
DR EMBL; AF339158; AAK38716.1;
DR EMBL; AF339159; AAK38717.1;
DR EMBL; AF339160; AAK38718.1;
DR EMBL; AF339161; AAK38719.1;
DR EMBL; AF339162; AAK38720.1;
DR EMBL; AF339163; AAK38721.1;
DR EMBL; AF339164; AAK38722.1;
DR EMBL; AF339165; AAK38723.1;
DR EMBL; AF339166; AAK38724.1;
DR EMBL; AF339167; AAK38725.1;
DR EMBL; AF339168; AAK38726.1;
DR EMBL; AF339169; AAK38727.1;
DR EMBL; AF339170; AAK38728.1;
DR EMBL; AF339171; AAK38729.1;
DR EMBL; AF339172; AAK38730.1;
DR EMBL; AF339173; AAK38731.1;
DR EMBL; AF339174; AAK38732.1;
DR EMBL; AF339175; AAK38733.1;
DR EMBL; AF339176; AAK38734.1;
DR EMBL; AF339177; AAK38735.1;
DR EMBL; AF339178; AAK38736.1;
DR EMBL; AF339179; AAK38737.1;
DR EMBL; AF339180; AAK38738.1;
DR EMBL; AF339181; AAK38739.1;
DR EMBL; AF339182; AAK38740.1;
DR EMBL; AF339183; AAK38741.1;
DR EMBL; AF339184; AAK38742.1;
DR EMBL; AF339185; AAK38743.1;
DR EMBL; AF339186; AAK38744.1;
DR EMBL; AF339187; AAK38745.1;
DR EMBL; AF339188; AAK38746.1;
DR EMBL; AF339189; AAK38747.1;
DR EMBL; AF339190; AAK38748.1;
DR EMBL; AF339191; AAK38749.1;
DR EMBL; AF339192; AAK38750.1;
DR EMBL; AF339193; AAK38751.1;
DR EMBL; AF339194; AAK38752.1;
DR EMBL; AF339195; AAK38753.1;
DR EMBL; AF339196; AAK38754.1;
DR EMBL; AF339197; AAK38755.1;
DR EMBL; AF339198; AAK38756.1;
DR EMBL; AF339199; AAK38757.1;
DR EMBL; AF339200; AAK38758.1;
DR EMBL; AF339201; AAK38759.1;
DR EMBL; AF339202; AAK38760.1;
DR EMBL; AF339203; AAK38761.1;
DR EMBL; AF339204; AAK38762.1;
DR EMBL; AF339205; AAK38763.1;
DR EMBL; AF339206; AAK38764.1;
DR EMBL; AF339207; AAK38765.1;
DR EMBL; AF339208; AAK38766.1;
DR EMBL; AF339209; AAK38767.1;
DR EMBL; AF339210; AAK38768.1;
DR EMBL; AF339211; AAK38769.1;
DR EMBL; AF339212; AAK38770.1;
DR EMBL; AF339213; AAK38771.1;
DR EMBL; AF339214; AAK38772.1;
DR EMBL; AF339215; AAK38773.1;
DR EMBL; AF339216; AAK38774.1;
DR EMBL; AF339217; AAK38775.1;
DR EMBL; AF339218; AAK38776.1;
DR EMBL; AF339219; AAK38777.1;
DR EMBL; AF339220; AAK38778.1;
DR EMBL; AF339221; AAK38779.1;
DR EMBL; AF339222; AAK38780.1;
DR EMBL; AF339223; AAK38781.1;
DR EMBL; AF339224; AAK38782.1;
DR EMBL; AF339225; AAK38783.1;
DR EMBL; AF339226; AAK38784.1;
DR EMBL; AF339227; AAK38785.1;
DR EMBL; AF339228; AAK38786.1;
DR EMBL; AF339229; AAK38787.1;
DR EMBL; AF339230; AAK38788.1;
DR EMBL; AF339231; AAK38789.1;
DR EMBL; AF339232; AAK38790.1;
DR EMBL; AF339233; AAK38791.1;
DR EMBL; AF339234; AAK38792.1;
DR EMBL; AF339235; AAK38793.1;
DR EMBL; AF339236; AAK38794.1;
DR EMBL; AF339237; AAK38795.1;
DR EMBL; AF339238; AAK38796.1;
DR EMBL; AF339239; AAK38797.1;
DR EMBL; AF339240; AAK38798.1;
DR EMBL; AF339241; AAK38799.1;
DR EMBL; AF339242; AAK38800.1;
DR EMBL; AF339243; AAK38801.1;
DR EMBL; AF339244; AAK38802.1;
DR EMBL; AF339245; AAK38803.1;
DR EMBL; AF339246; AAK38804.1;
DR EMBL; AF339247; AAK38805.1;
DR EMBL; AF339248; AAK38806.1;
DR EMBL; AF339249; AAK38807.1;
DR EMBL; AF339250; AAK38808.1;
DR EMBL; AF339251; AAK38809.1;
DR EMBL; AF339252; AAK38810.1;
DR EMBL; AF339253; AAK38811.1;
DR EMBL; AF339254; AAK38812.1;
DR EMBL; AF339255; AAK38813.

30-MAY-2000 (Rel. 39, Created)
 30-MAY-2000 (Rel. 39, Last sequence update)
 15-JUN-2002 (Rel. 41, Last annotation update)
 Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 protease 2) (ASP2) (Membrane-associated aspartic protease 2)
 (Memapsin-2).
 GN BACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A., Luo Y.,
 Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 Treanor J., Rogers G., Citron M.;
 "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
 the transmembrane aspartic protease BACE.";
 Science 286:735-741(1999).
 RL -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
 CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
 CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
 CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
 CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
 CC FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF190727; AAF04144.1; -
 CC HSP: P33229; IYPS.
 CC MERO: A01.004; -
 CC InterPro: IPR001461; AsparticaseA1.
 CC InterPro: IPR001969; AsparticaseA1.
 CC Pfam: PF00026; asp. 1.
 CC PRINTS: PF00792; PEPSIN.
 CC PROSITE: PS00141; ASP_PROTEASE; 1.
 KW Hydrolyase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 KW Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 45 POTENTIAL.
 FT CHAIN 46 501 BETA-SECRETASE.
 FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 FT DISULFID 216 420 BY SIMILARITY.
 FT DISULFID 278 443 BY SIMILARITY.
 FT DISULFID 330 380 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 501 AA; 55806 MW; 24B445BC8BE87DE3 CRC64;
 Query Match 96.4%; Score 2569; DB 1; Length 501;
 Best Local Similarity 96.2%; Pred. No. 8.2e-199;
 Matches 482; Conservative 12; Indels 0; Gaps 0;
 Oy 1 MAQALPWLMMGAGVLPAGHTQHGIRLPLRSLGGAPLGLRLPRETDEPEPGRGSF 60

1 MAPALRWLLWVSGMLPAQOGLHGLRPLRSLGAPLGLRLPRETDEPEPGRGSF 60
 61 VEMVDNLGRSGGQYVEMTVGSPPTNLILVDGSSNFVAGAAPHPFLHRYQRLSST 120
 61 VEMVDNLGRSGGQYVEMTVGSPPTNLILVDGSSNFVAGAAPHPFLHRYQRLSST 120
 121 YRDLRKGVYPTQGWEGELGTLVSTPHGPNVTVRANITAAITSDKFFINGSNWEGIL 180
 121 YRDLRKSVYPTQGWEGELGTLVSTPHGPNVTVRANITAAITSDKFFINGSNWEGIL 180
 181 GLAYAEIARPDSDLPEPFDSLVKQTHVPNLSFLHLCGAGPFLNQSEVLASVGGSMIIGGI 240
 181 GLAYAEIARPDSDLPEPFDSLVKQTHVPNLSFLHLCGAGPFLNQSEVLASVGGSMIIGGI 240
 241 DHSLYTGLSWYTPIRREMYEYVIRVEINQODLKMDCKEYNDKSIDVSGTTNLRPKK 300
 241 DHSLYTGLSWYTPIRREMYEYVIRVEINQODLKMDCKEYNDKSIDVSGTTNLRPKK 300
 301 VFEAAVKSIRAASTEFKPDGFWLGEOLVQWAGCTTPWNIFFPVISLYLMGEVNTNOSFRIT 360
 301 VFEAAVKSIRAASTEFKPDGFWLGEOLVQWAGCTTPWNIFFPVISLYLMGEVNTNOSFRIT 360
 361 ILPOQYLRPVEDVATSDQCYKFAISQSTGTVMGAVIMEGFYVDFRARRKRGFVVSAC 420
 361 ILPOQYLRPVEDVATSDQCYKFAISQSTGTVMGAVIMEGFYVDFRARRKRGFVVSAC 420
 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDSTLMTIAYVMAAICAFMLPLCLMVCOW 480
 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQDSTLMTIAYVMAAICAFMLPLCLMVCOW 480
 481 RCLRLRQHQHDDFADDSLLK 501
 481 RCLRLRQHQHDDFADDSLLK 501
 RESULT 3
 BACE_MOUSE
 ID BACE_MOUSE STANDARD; PRG; 501 AA.
 AC P56818;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
 DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
 DE protease 2) (ASP2) (Membrane-associated aspartic protease 2)
 DE (Memapsin-2).
 DE BACE.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 Treanor J., Rogers G., Citron M.;
 "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
 the transmembrane aspartic protease BACE.";
 Science 286:735-741(1999).
 RL [2]
 RP REVISIONS TO 6 AND 81-87.
 RA Bennett B.D., Vassar R., Citron M.;
 Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20051710; PubMed=10591213;
 RA Yan R., Bienkowska M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
 Brashler J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
 Tomasselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.E.;

RT RT "Membrane-anchored aspartyl protease with Alzheimer's disease
RL Nature 402:533-537(1999).
CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC THE A-BETA PEPTIDE SEQUENCE. BETWEEN RESIDUES 671 AND 672 OF APP.
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF190726; AAF04143.2; -
CC EMBL: AF200346; AAF17082.1; -
CC HSP: P56272; IAM5.
CC MERO: A01.004; -
CC MGD; MG1:1346542; Bace.
CC InterPro: IPR001461; Asparticase1.
CC InterPro: IPR001969; Asparticase_site.
CC Pfam: PF00026; asp; 1.
CC PRINTS: PR00792; PEPsin.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC Hydrolase; Aspartyl protease; Zymogen; Transmembrane;
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 45 POTENTIAL.
FT CHAIN 46 501 BETA-SECRETASE.
FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 458 478 POTENTIAL.
FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 289 289 BY SIMILARITY.
FT DISULFID 216 420 BY SIMILARITY.
FT DISULFID 278 443 BY SIMILARITY.
FT DISULFID 330 380 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 501 AA; 55747 MW; C085A013145E47AE CRC64;
Query Match 96.4%; Score 2567; DB 1; Length 501;
Best Local Similarity 96.2%; Pred. No. 1.2e-198;
Matches 482; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
Qy 1 MAQALPWLWVGALVPAHQTHGIRLPLRSLGGLGAPLGLRLPRETDEPEGRRSF 60
Db 1 MAPALHLLWVGSGMLPAQCTHLGIRLPLRSLGAGPLGLRLPRETDEPEGRRSF 60
Qy 61 VEWVDNLRGSGGQYVYVEMTVGSPPTLNLVDTGSSNFAVGAAPFHLRYQRLSST 120
Db 61 VEWVDNLRGSGGQYVYVEMTVGSPPTLNLVDTGSSNFAVGAAPFHLRYQRLSST 120
Qy 121 YRLRKGVYVPYQGWKGEGLDLSVPHGPNVTVRANTAATESDKFFNGSNWEGIL 180
Db 121 YRLRKGVYVPYQGWKGEGLDLSVPHGPNVTVRANTAATESDKFFNGSNWEGIL 180
Qy 181 GLAYAEIARDDSLPEPFDLSLVKQTHVPNLFSLHLCGAGFPNQSEVLASVGGSMIIIGI 240
Db 181 GLAYAEIARDDSLPEPFDLSLVKQTHVPNLFSLHLCGAGFPNQSEVLASVGGSMIIIGI 240
Qy 241 DHSLYTSLWYTPIRREWEYVEIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Db 241 DHSLYTSLWYTPIRREWEYVEIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300

Db 241 DHSLYTSLWYTPIRREWEYVEIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Qy 301 VFEAAVKSIIKAASSTKEKPDGFWLGEOLVQVQAGTTPWNIPFVLSLYLMGEVTNOSFRIT 360
Db 301 VFEAAVKSIIKAASSTKEKPDGFWLGEOLVQVQAGTTPWNIPFVLSLYLMGEVTNOSFRIT 360
Qy 361 ILPQOYLPRVEDVATSDQCYKFAISQSGSTGTGAVGAVIMEGFYVDFRARRKRGFAVSAC 420
Db 361 ILPQOYLPRVEDVATSDQCYKFAISQSGSTGTGAVGAVIMEGFYVDFRARRKRGFAVSAC 420
Qy 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
Db 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
Qy 481 RCLRLRQHQHDDFADDISLLK 501
Db 481 RCLRLRQHQHDDFADDISLLK 501
RESULT 4
ID_BAE2_HUMAN STANDARD: PRT: 518 AA.
AC Q9Y5Z0; Q9JUT6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.-) (Beta-site APP-cleaving
DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASPI) (Membrane-associated
DE aspartic protease 1) (Memapsin-1).
GN BACE2 OR ASP21
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT "Membrane-anchored aspartyl protease with Alzheimer's disease
RT beta-secretase activity.";
RL Nature 402:533-537(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
RA Giese K.;
RT "Identification of a novel aspartic-like protease differentially
RT expressed in human breast cancer cell lines.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
RT transmembrane aspartyl protease.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Solans A., Estivill X., de la Luna S.;
RT "Cloning of a novel mammalian aspartyl protease.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20120043; PubMed=10656250;
RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
RT "Identification of a novel aspartic proteinase (Asp 2) as
RT beta-secretase.";
RL Mol. Cell. Neurosci. 14:419-427(1999).
RN [6]
RP SEQUENCE FROM N.A.

Matches	239;	Conservative	82;	Mismatches	165;	Indels	33;	Gaps									
QY	2	AQALPWL	LWM---	GAGVLP	PAHQHG	IRLP	LRSLG	-----GAPL-----GLR 42									
Db	7	ALLLP	LAQWLL	RAAP	ELAP	APPT----	LPLR	VAAATNRVAPT									
QY	43	LPR	TEDEPE	PEPGR	GRSGV	EMVDN	LRKSG	GOGYVYEMT									
Db	62	LAL	E--PAL	ASPA	GAAN	FLAMVDN	LQD	SGRGYVLEMLIGT									
QY	103	RA	PF	LHRYQ	R	QLSS	YRDLR	KGVYVYPTQGW									
Db	120	GP	PHSY	IDTY	EDT	ERS	YSR	SGFDVTKYQTQGS									
QY	163	IT	ESK	FFINGS	NW	EGIL	GLAY	AEIAR									
Db	180	IF	EN	FFEL	PG	IKWNG	ILGL	AYATLAK									
QY	223	NO	SEVL	ASV	GGSM	ITGG	IDH	SLYTGSLWYTP									
Db	240	AG	S--GT	NGS	VL	GGI	EP	SLYKGD									
QY	283	YD	K	SV	D	SCT	TN	LRLPK									
Db	297	AD	K	AI	V	D	SCT	TN									
QY	343	V	T	S	L	M	G	E									
Db	357	K	I	S	I	V	R	D									
QY	403	Y	V	F	D	R	A	K									
Db	416	Y	V	F	D	R	A	K									
QY	463	A	A	I	C	-A	L	F									
Db	476	M	S	V	G	A	I	L									
RESULT 5																	
ID	PEPI_GADMO	STANDARD;	PRT;	324	AA.												
AC	P56272;																
DT	15-JUL-1998	(Rel. 36, Created)															
DT	15-JUL-1998	(Rel. 36, Last sequence update)															
DT	15-JUN-2002	(Rel. 41, Last annotation update)															
DE	Pepsin IIB (EC 3.4.23.-).																
OS	Gadus morhua (Atlantic cod).																
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;																
OC	Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.																
OX	NCBI_TaxID=8049;																
RN	[1]																
SEQUENCE, AND X-RAY CRYSTALLOGRAPHY.																	
RP	TISSUE=Stomach;																
RA	Karlisen S., Hough E., Olsen R.L.;																
RC	"Structure and proposed amino-acid sequence of a pepsin from Atlantic																
RT	cod (Gadus morhua)".																
RT	Acta Crystallogr. D 54:32-46(1998).																
RL	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.																
CC	PDB; 1AM5; 24-DEC-97.																
DR	InterPro; IPR001461; AspproteaseA1.																
DR	InterPro; IPR001969; Aspprotease_site.																
DR	Pfam; PF000026; asp; 1.																
DR	PRINTS; PRO0792; PEPSIN.																
DR	PROSITE; PS00141; ASP_PROTEASE; 2.																
KW	Hydrolase; Aspartyl protease; Digestion; 3D-structure.																
FT	ACT_SITE	32	32	BY SIMILARITY.													
FT	ACT_SITE	214	214	BY SIMILARITY.													
FT	DISULFID	45	50	BY SIMILARITY.													

Query Match	44.18	Score 1173.5; DB 1; Length 518;
Best Local Similarity	46.18	Pred. No. 1.2e-86;

```

Query Match      12.4%; Score 330; DB 1; Length 324;
Best Local Similarity 27.9%; Pred. No. 3.7e-19;
Matches 104; Conservative 67; Mismatches 136; Indels 66; Gaps 15;

QY 63 MVDNLKSGQGYVEMTGVSPQTLNIIIVDTGSSNFVAG----AAPHFLHRYVORLS 118
DB 2 VTEQMKNEADTEYGVISIGTTPESFKVIFDTGSSNLWVSSSHCSAQACSNHNKFKPQS 61

QY 119 STYDLRKGVVYDTGKWEGLGTLDSIPHG--PNVTVRANIAAITESDKFEFNGSNW 176
DB 62 STYVETGKTVDLRYGCGMRGILGQDVTSGVGGSPNOELG---ESQTEPGFQA-AAPF 117

QY 177 EGTGLGAYAIAPDDSLPEFDSLKVQTHV--PNFLSLHLCAGFPPLNOSSEVLASVGGSM 235
DB 118 DGILGLAYPSIAAA--GAYPVFDNMGSSQILVEKDLFFSYLGGG--ANGSEVM----- 166

QY 236 IIGIDHSLYTGSLWTPYPIREYVYVIVRVEINGQDLKMD--CKEYNDKSIDVSGTGN 294
DB 167 -LGVDSNHYTGSIHWPVTAERYQWALDGTITVNGQTAAECGC-----QALVDGTGSK 219

QY 295 LRLPKKVFEEAAVKSIAASSTKFPDGFGLQVLCVQWAGTTPWNIFFPVISLYLMGEVTN 354
DB 220 IVAPVSALANIMKDIGASEN-----OGEMGN--CASVGSPLDITF-----TI 260

QY 355 QSFRTITLPOOYLPRVEDVATSDODCYKFAISOSSSTGT-----VMGAVIMEGFYVVF 406
DB 261 NGVKOPLPPSAYIEGQAFCTSD-----GLGSGVPSNTSELWIFGDVFLURNYITTY 311

QY 407 DRARKRIGFAVSA 419
DB 312 DRTNKKVGFAPAA 324

RESULT 6
CATD_BOVIN STANDARD; PRT; 390 AA.
AC AC P80209; Q9TS27;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
SEQUENCE OF 1-48.
TISSUE=Milk;
MEDLINE=93202276; PubMed=8454061;
RA Larsen L.B., Boisen A., Petersen T.E.;
RT "Procathepsin D cannot autoactivate to cathepsin D at acid pH.";
RL FEBS Lett. 319:54-58(1993).
[2]
SEQUENCE OF 45-390, AND X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
RT "Two crystal structures for cathepsin D: the lysosomal targeting
signal and active site.";
RL EMBO J. 12:1293-1302(1993).
CC CC -1- FUNCTION: Acid protease active in intracellular protein breakdown.
CC -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC that of pepsin A. Does not cleave the 4-Gln-His-5 bond in B
chain of insulin.
CC -1- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR PIR: S32383; S32383.
DR PIR: S37419; S37419.
DR HSSP: P07339; 1LYB.

```

```

MEROPS; A01.009;
DR InterPro: IPR001461; AsparticaseA1.
DR InterPro: IPR001969; Asparticase_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00192; PEPSIN.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen.
FT PROPEP 1 44
FT CHAIN 45 390
FT ACT_SITE 77 77
FT ACT_SITE 273 273
FT DISULFID 71 140
FT DISULFID 90 97
FT DISULFID 264 268
FT DISULFID 307 344
FT CARBOHYD 114 114
FT CARBOHYD 241 241
SQ SEQUENCE 390 AA; 42488 MW; 5B38AAIC33C48D35 CRC64;

Query Match      11.8%; Score 314.5; DB 1; Length 390;
Best Local Similarity 28.0%; Pred. No. 8.4e-18;
Matches 113; Conservative 72; Mismatches 128; Indels 91; Gaps 21;

QY 53 EPG-RRGSFVEMVDNLRCKSGQGYVEMTGVSPQTLNIIIVDTGSSNFVAGAAHPHPL-- 109
DB 39 EPAVRGQPIPELLKNYMDAQ---YGEIGIGTTPQCFVVDFTGSANLWVPSIHKLLDI 95

QY 110 ---HRYQRLSSTYRLRKGVY--VPYTGKWEGLGTLDSIPHGPN-----VTVR 157
DB 96 ACWTHRYKNSDKSSTY--VKNGTTFDIHYGSGLSGYLSQDTSVPCNPSSSSPGGVTVQ 153

QY 158 ANI--AAITESDKFFINGSNWEGILGLAYAIAPDDSLPEFDSLKVQTHV--PNFLSLH 214
DB 154 RQTGGEAIKQGVVFI-AAKFDGILGMAYPRIS--VNNVLVFDNLMQOKLVKNVFS-- 208

QY 215 LCGAGFPLNOSSEVLASVCGSMIIGIDHSLYTGSLWTPYPIREYVYVIVRVEINGQDL 274
DB 209 -----FFLNR-DPKAQPGGELMGLGTDKSKYRGSLFHNVTROAYWQHMDQLDV-GSSL 261

QY 275 KMDCKEYNYDKSIDVSGTTLNLRPKKVFEEAAVKSIAASSTKFPDGFGLQVLCVQW 333
DB 262 TV-CK--GGCEAIVDTGTSLVGVPVEEVRLOKAIGAVPLIQ-----GEYIIPCEKV 310

QY 334 GTTPWNIFFVISLYLMGEVTNQSFRITLPOOYLPRVEDVATSDODCYKFAISOSSSTGT- 392
DB 311 SS-----LPEVVKLG-----KDYALSPED-YALKVSOAETVVC 344

QY 393 -----VMGAVIMEGFYVVFDRARKRIGFAVSA 419
DB 345 LSGFMGMIDIPPPGGPLWILGDFVIGRYTVFDRDQNRVGLAEAA 388

RESULT 7
PEPI_RABIT STANDARD; PRT; 387 AA.
ID PEPI_RABIT
AC P28712;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pepsin II-1 precursor (EC 3.4.23.1) (Pepsin A).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=91009127; PubMed=2129536;
RA Kageyama T., Tanabe K., Koizumi O.;
RT "Structure and development of rabbit pepsinogens. Stage-specific
zymogens, nucleotide sequences of cDNAs, molecular evolution, and
gene expression during development.";
RL J. Biol. Chem. 265:17031-17038(1990).
CC -1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS

```


CC	INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
CC	ALSO CLEAVED TO SOME EXTENT.
CC	-I- CATALYTIC ACTIVITY: preferential cleavage: hydrophobic, preferably
CC	aromatic, residues in p1 and p1' positions. Cleaves 1-Phe-1-Val-2,
CC	4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Leu-15, 15-Leu-1-Tyr-16,"
CC	16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-
CC	Tyr-26 bonds in the B chain of insulin.
CC	-I- DEVELOPMENTAL STAGE: PEPSINOGENS IN GROUP I, II, AND III WHERE
CC	THE PREDOMINANT ZYMOCENS AT LATE POSTNATAL STAGE.
CC	-I- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC	HORMONES AND RELATED SUBSTANCES.
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR	PIR; B38302; B38302.
DR	HSSP; P00791; IPSA.
DR	MEROPS; A01.001; -.
DR	InterPro; IPR001461; AsparticaseA1.
DR	InterPro; IPR001969; Asparticase_site.
DR	Pfam; PF00026; asp; 1.
CC	PRINTS; PR00792; PEPsin.
CC	PROSITE; PS00141; ASP_PROTEASE; 2.
CC	Hydrolase; Aspartyl protease; Digestion; Zymogen; Signal;
CC	Phosphorylation; Multigene family.
KW	SIGNAL 1 15
FT	PROPEP 16 59 ACTIVATION PEPTIDE.
FT	CHAIN 60 387 PEPsin II-1.
FT	MOD_RES 129 129 PHOSPHORYLATION (POTENTIAL).
FT	ACT_SITE 93 93 BY SIMILARITY.
FT	ACT_SITE 276 276 BY SIMILARITY.
FT	DISULFID 106 111 BY SIMILARITY.
FT	DISULFID 267 271 BY SIMILARITY.
FT	DISULFID 310 343 BY SIMILARITY.
SQ	SEQUENCE 387 AA; A6EC48F715541A48 CRC64;
Query Match	11.58; Score 309; DB 1; Length 387;
Best Local Similarity	27.18; Pred. No. 2.3e-17;
Matches	98; Conservative 68; Mismatches 130; Indels 66; Gaps 15;
QY	75 YVYEMTVGSPQPTNLILVDTGSSNFVAG---AAPHFPLHRYXORQLSSYVRLRKGVYV 130
DB	75 YFTGISLTPQPEFTVIFDYGSSNLWVPSTYCSLACLHRRNPDDSSFTQATSETLSI 134
QY	131 PYTQGWKEGLGDLVSIHPGPNVTVRANTAATESD---KFFINGSNWEGILGLAYAEI 187
DB	135 TYGTGSMTGILGYDVKV---GNIEDNQIFGLSKTEPGITFLV--APFDGILGLAYPSI 189
QY	188 ARPDDSLPEPFDLSLVKQTHV--PNLFSUHLGCAGFPLNQSEVLASVGSMITGGIDHSLYT 246
QY	190 SASDAT--PVFDNNWNEGLVSEDLFSVYLSNG-----EKGSMWFGGIDSSVYT 237
QY	247 GSLWYPIRREWYEVIIIVRVEINGDQLKM--DCKEYNDKSIYDSTGTTNLRLPKVFEEA 304
DB	238 GSNWVPSVSEHGVIQMTDSITTINGETIACADSC-----QAVYDTCSTLLAGTSAISK 291
QY	305 AVKSIAAASTEKFPDGFWLGEOLV--CWQAGTTPWNIFPVISLYLMGEVTVNQSFRTILP 363
DB	292 IQSYIGASKNL-----LGENIISCAISDLSLDIVF-----TINN 325
QY	364 QOYLRLPYED--VATSQDDC---YKFAISQSSTGT--VMGAVIMEGVYVVFDRKRIGFAV 417
DB	326 VQYPLPASAYILKEDDDCLSGFDGMNLDTSYGELWILGVDVFIQYFTVFDPRANNQVGLAA 385
QY	418 SA 419
DB	386 AA 387

RESULT 8		PEP4_MACFU		STANDARD;		PRT: 388 AA.	
ID	PEP4_MACFU						
AC	P27678;						
DT	01-AUG-1992	(Rel. 23, Created)					
DT	01-AUG-1992	(Rel. 23, Last sequence update)					
DT	15-JUN-2002	(Rel. 41, Last annotation update)					

Pepsin A-4 precursor (EC 3.4.23.1) (Pepsin I/II).
PCA.
Macaca fuscata fuscata (Japanese macaque).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
NCBI_TaxID=9543;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 16-70.
TISSUE=Gastric mucosa;
MEDLINE=92037645; PubMed=1935977;
Kageyama T., Tanabe K., Koizumi O.;
"development-dependent expression of isozymogens of monkey
pepsinogens and structural differences between them.";
Eur. J. Biochem. 202;205-215(1991).
-1- FUNCTION: SHOWS PARTICULARLY BROAD SPECIFICITY; ALTHOUGH BONDS
INVOLVING PHENYLALANINE AND LEUCINE ARE PREFERRED, MANY OTHERS ARE
ALSO CLEAVED TO SOME EXTENT.
-1- CATALYTIC ACTIVITY: Preferential cleavage: hydrophobic, preferably
aromatic, residues in P1 and P1' positions. Cleaves 1-Phe-1-Val-2,
4-Gln-1-His-5, 13-Glu-1-Ala-14, 14-Ala-1-Leu-15, 15-Leu-1-Tyr-16,
16-Tyr-1-Leu-17, 23-Gly-1-Phe-24, 24-Phe-1-Phe-25 and 25-Phe-1-
Tyr-26 bonds in the B chain of insulin.
-1- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
HORMONES AND RELATED SUBSTANCES.
-1- PEPSINOGENS: EACH PEPSINOGEN IS CONVERTED TO CORRESPONDING
PEPSIN AT PH 2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA
ACTIVATION SEGMENT AND IN PART AS A RESULT OF STEPWISE PROTEOLYTIC
CLEAVAGE VIA AN INTERMEDIATE FORM(S).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; X59753; CAA42425.1; --
PIR; S16065; S16065.
PIR; S19682; S19682.
HSP; P00790; IPSN.
MEROPS; A01.001; --
InterPro: IPR001461; AsproteaseA1.
InterPro: IPR001969; Asprotease_site.
InterPro: IPR001791; Laminin_G.
Pfam: PF00026; asp_1
PRINTS; PR00792; PEPSIN.
PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Zymogen; Multigene family;
KW Signal; Glycoprotein.
FT SIGNAL 1 15 BY SIMILARITY.
FT PROPEP 16 38 ACTIVATION PEPTIDE.
FT PROPEP 39 62 ACTIVATION PEPTIDE.
FT CHAIN 63 388 PEPSIN A-4.
FT ACT_SITE 94 94 BY SIMILARITY.
FT ACT_SITE 277 277 BY SIMILARITY.
FT DISULFID 107 112 BY SIMILARITY.
FT DISULFID 268 272 BY SIMILARITY.
FT DISULFID 311 344 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC...).
SQ SEQUENCE 388 AA; A1955 MW; A2923AB1F7CDEB9 CRC64;

Query Match	11.5%	Score	307.5;	DB	1;	Length	388;
Best Local Similarity	27.6%	Pred. No.	3.1e-17;				
Matches	108;	Conservative	65;	Mismatches	135;	Indels	83;
						Gaps	17;
Qy	44	PRETDEEPEPGRGSGFVEMVDNLKSGOGYVEMTVGSPPTLTNLVDTGTSNFAVGA	103				
Db	60	PTLIDEQPLE----	NYLDD-----	EYETIGIGTPAQFTVFVDTGTSNLWV--	102		
Qy	104	APHPFL-----	HRYYQRLSTYRLDRKGVVVPYTGQKWEGLGDLTVSIPHGPNVT	156			

103	DB	-PSVYCSLACMDHNLNPODSTYRTSKTVSITVGTGSMTGILGYDVKV----	GGISD	158
157	QY	RANIAAATESDK--FFINGSNWEGILGLAYAEIARPDLSLEPPFDSLVKQTHV--PNLFLSLH	214	
159	DB	TNQIFGLSETPGFFLYFAPDGLGLAYPSIS--SSGATPVFDNINWQRLVSQDLFSYI	216	
215	QY	LCGAGFPNLQSEVLASVGGSMIGIDHSLYTGSLWTPIRREWEYVEIIVRVEINGQDL	274	
217	DB	LSAD-----DQS-----GSVIFGIDSSYTGSLNWPVSVEGYWQISVDSITMNGKI	266	
275	QY	--KMDCKEYNDKSGIVSGTNNLRPKKVEAAVKSIAAASSTKPPDGFVLGEOLV-CW	331	
267	DB	ACAKGC-----QAIVDTGTSLLAGTSPIANIQSDICASENSD-----GEMVVSCS	312	
332	QY	QAGTTPWNIFFVISLYLMGEVNTNSFRITILPOQY-LRPVEDVATSDDCYK-----FAI	385	
313	DB	AISSLPLDIVF-----TINGVOYPLPPSAIILQSQSCTSGFGQMDVP	354	
386	QY	SQSTGTVMGAVIMEGFFVWFDRAKRIGFA	416	
355	DB	TESGELWILGDVFIQYFTVFDRAANNQVGLA	385	

103	DB	YVEMTVGSPQDTNLIIVDTGSSNFAYGAAPHPPL-----HRYORQLSSTYRDLRKG	127	
59	DB	YVGTISITGTPQODFSVIFDTGSSNLWV---PSIYCKSSACSNHKKRDPKSSSIYVSTNET	115	
128	QY	VVPVYTGCKWGEGLGTLVSIPIHPGNVNTVRANIAAATESDK--FFINGSNWEGILGLAYAE	186	
116	DB	VVIAGTGSMSGLIGYDIVAV---SSIDVQNOIFGLSETEPGSPFFYCNFDGILGLAFPS	172	
187	QY	IARPDLSLEPPFDSLVKQTHV--PNLFLSLHLCGAGFPNLQSEVLASVGGSMITGIDHSLY	245	
173	DB	IS--SSGATPVFDNINWQRLVSQDLFSVLSKDG-----ETGSFVLFGIDFNIT	220	
246	QY	TGSLWTPIRREWEYVEIIVRVEINGQDLK--MDCKEYNDKSIDVSGTNNLRPKKVE	303	
221	DB	TKGIWVPLSAETVWQITMDRVTVGNKYVACFFTC-----QAIVDTGTSLLVMPQAGYN	274	
304	QY	AAVKSIAASSTK-----KPDGFWLGEOLVQWAGTTPWNIFFVISLYLMGEVNTNS	356	
275	DB	RIIKDLGVSSDGEISCDSDISKLPD-----VTFHINGHA-----	307	
357	QY	FRTILPOQYLRPVEDVATSDDCYKFAISQSTGT-----VMGAVIMEGFFVWFDRA	410	
308	DB	FTLPASAYVLNEDGSCMLGFENNGTPTGELGEQWILGDVFIQYFTVFDRA	358	
411	QY	KRIGFA	416	
359	DB	NKVGLS	364	

AC	PI6476;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Embryonic pepsinogen precursor (EC 3.4.23.-)			
OS	Gallus gallus (Chicken)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88227903; PubMed=3131317;			
RA	Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;			
RT	"Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prothymosin.";			
RL	J. Biochem. 103:290-296(1988).			
CC	DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).			
CC	!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; D00215; BAA00153.1;			
DR	PIR; A41443; A41443.			
DR	HSSP; P00794; 4CMS.			
DR	MEROPS; A01.028;			
DR	InterPro; IPR001461; AspproteaseA1.			
DR	InterPro; IPR001969; Aspprotease_site.			
DR	Pfam; PF00026; asp; 1.			
DR	PRINTS; PR00792; PEPSIN.			
DR	PROSITE; PS00141; ASP_PROTEASE; 2.			
KW	Hydrolase; Aspartyl protease; Digestion; Signal; Glycoprotein.			

AC	PI6476;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Embryonic pepsinogen precursor (EC 3.4.23.-)			
OS	Gallus gallus (Chicken)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88227903; PubMed=3131317;			
RA	Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;			
RT	"Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prothymosin.";			
RL	J. Biochem. 103:290-296(1988).			
CC	DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).			
CC	!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; D00215; BAA00153.1;			
DR	PIR; A41443; A41443.			
DR	HSSP; P00794; 4CMS.			
DR	MEROPS; A01.028;			
DR	InterPro; IPR001461; AspproteaseA1.			
DR	InterPro; IPR001969; Aspprotease_site.			
DR	Pfam; PF00026; asp; 1.			
DR	PRINTS; PR00792; PEPSIN.			
DR	PROSITE; PS00141; ASP_PROTEASE; 2.			
KW	Hydrolase; Aspartyl protease; Digestion; Signal; Glycoprotein.			

FT	CHAIN	43	367	
FT	ACT_SITE	77	77	
FT	ACT_SITE	260	260	
FT	CARBOHYD	113	113	
FT	DISULFID	90	95	
FT	DISULFID	251	255	
FT	DISULFID	290	323	
SQ	SEQUENCE	367 AA;	40431 MW;	OC547E7FD8F5B341 CRC64;

Query Match	11.4%;	Score 305;	DB 1;	Length 367;
Best Local Similarity	24.0%;	Pred. No. 4.5e-17;		

Db 111 SPACKTHSRQPSQSTYSQPGSQSFQYGTGSLGIGADQVSV-EGLTVVQOQFGESV 169

Qy 164 TESDKFTINGSNWEGLILGAYAEIARPDSDLEPFDSLVKQTHVP-NLFSHLCCGAGFL 222

Db 170 TEPQGTVD-AEFDGLGLGYPSLA--VGGVTPVFDNMAQNLVDLPMSFVY- 219

Qy 223 NOSEVLASVSGSMIIGDHSHTGSLWYTPIPREWYEVIIYVRVEINGODLKMOCKEYN 282

Db 220 -SSNPEGAGSELIFGVDHSHFSGSLNWVPVTKQAYQWQIALDNIQVGG-TVMFCSE-- 274

Qy 283 YDKSIVSGTNNLRPKKVFEEAAVKSAAASSTKEKFPDGFNLGELQVLCWQAGTTPWNIPP 342

Db 275 GCOAIVDTGTSLITGSPDKIKQLQNAIGAAP-----VDGEYAVE-----CANLVMP 321

Qy 343 VISLYLMEVNTQSFRTILPQOYLRPVEDVATSDQDCYKFAISQSTG----- 391

Db 322 DVTFTING-----VPYTLSTAY--TLLDPVDMQF-----SSGFGQLDIHPAG 365

Qy 392 --TVMGAVIMEGYVVFEDRARKRIGFA 416

366 PLWILGDFVIRQYFVDFGRNVRVGLA 392

RESULT 12

ID	CATD_HUMAN	STANDARD;	PRT;	412 AA.
AC	P07339;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Cathepsin D precursor (EC 3.4.23.5).			
GN	CTSD.			
OS	"Homo sapiens (human)."			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]	SEQUENCE FROM N.A.			
RP	MEDLINE=85270436; PubMed=3927292;			
RX	Faust P.L., Kornfeld S., Chirgwin J.M.;			
RA	"Cloning and sequence analysis of cDNA for human cathepsin D.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).			
[2]	SEQUENCE FROM N.A.			
RP	MEDLINE=87231058; PubMed=3598310;			
RX	Westley B.R., May F.E.B.;			
RA	"Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive			
RT	human breast cancer cells.;"			
RL	Nucleic Acids Res. 15:3773-3786(1987).			
[3]	SEQUENCE FROM N.A.			
RP	MEDLINE=91299159; PubMed=2069717;			
RX	Redecker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;			
RA	"Molecular organization of the human cathepsin D gene.;"			
RL	DNA Cell Biol. 10:423-431(1991).			
[4]	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
[5]	SEQUENCE OF 1-22 FROM N.A.			
RP	MEDLINE=94085791; PubMed=8262386;			
RX	May F.E., Smith D.J., Westley B.R.;			
RA	"The human cathepsin D-encoding gene is transcribed from an estrogen-			
RT	regulated and a constitutive start point.;"			
RL	Gene 134:277-282(1993).			
[6]	SEQUENCE OF 1-22 FROM N.A.			
RP	MEDLINE=95021301; PubMed=7935485;			
RX	Augereau P., Miralles F., Cavailles V., Gaudelet C., Parker M.;			
RA	Rochefort H.;			

"Characterization of the proximal estrogen-responsive element of human cathepsin D gene.;"

Mol. Endocrinol. 8:693-703(1994).

[7]

SEQUENCE OF 170-180.

TISSUE=Liver;

RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,

RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,

RA Appel R.D., Hughes G.J.;

RL Submitted (JUN-1992) to the SWISS-PROT data bank.

[8]

VARIANT VAL-58.

RP MEDLINE=20179010; PubMed=10716266;

RX Papassotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,

RA Maier W., Pauls J., Lautenschlager N., Heun R.;

RA "A genetic variation of cathepsin D is a major risk factor for

RT Alzheimer's disease.;"

RL Ann. Neurol. 47:399-403(2000).

[9]

X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).

RP TISSUE=Spleen;

RC MEDLINE=93223670; PubMed=8467789;

RA Metcalf P., Fusek M.;

RA "Two crystal structures for cathepsin D: the lysosomal targeting

RT signal and active site.;"

RL EMBO J. 12:1293-1302(1993).

[10]

X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).

RP TISSUE=Liver;

RC MEDLINE=93342076; PubMed=8393577;

RX Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,

RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;

RA "Crystal structures of native and inhibited forms of human cathepsin

RT D: implications for lysosomal targeting and drug design.;"

RL Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).

CC -i- FUNCTION: Acid protease active in intracellular protein breakdown.

CC Involved in the pathogenesis of several diseases such as breast

CC cancer and possibly Alzheimer's disease.

CC -i- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,

CC that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B

CC chain of insulin.

CC -i- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.

CC -i- SUBCELLULAR LOCATION: Lysosomal.

CC -i- POLYMORPHISM: The Val-58 allele is significantly overrepresented

CC in demented patients (11.8%) compared with nondemented controls

CC (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased

CC risk for developing AD than noncarriers.

CC -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC -----

CC EMBL; M11233; AAB59529.1; -

CC EMBL; X05344; CA28955.1; -

CC EMBL; M63138; AAA51922.1; -

CC EMBL; M63134; AAA51922.1; JOINED.

CC EMBL; M63135; AAA51922.1; JOINED.

CC EMBL; M63136; AAA51922.1; JOINED.

CC EMBL; M63137; AAA51922.1; JOINED.

CC EMBL; BC016320; AAL16320.1; -

CC EMBL; L12980; AAL16314.1; -

CC EMBL; S74689; AAD14156.1; -

CC EMBL; S52557; AAD13868.1; -

CC PIR; A25771; KKHUO.

CC PDB; 1LYA; 31-JAN-94.

CC PDB; 1LYB; 31-JAN-94.

CC MEROPS; A01.009; -

CC SWISS-2DPAGE; P07339; HUMAN.


```

Qy 354 NQSFRTILPQQYLKRPVEDVATSDCCYKFAISOSTGT-----VNGAVIMEGFYVVD 407
Db 322 -----TINGIQYPLPASAYILKEDDCTSGLEGMMVDTYTGTGELWILGDVFIQRYFTVD 375
Qy 408 RARRKRIIGFAVS 418
Db 376 RANNQLGLAAA 386

RESULT 15
CATD_RAT CATD_RAT STANDARD; PRT; 407 AA.
AC P24268;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
CN CTSD.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE=Pituitary;
MEDLINE=910571150; PubMed=2243802;
RA Birch N.P., Loh Y.P.;
RT "Cloning, sequence and expression of rat cathepsin D.";
RL Nucleic Acids Res. 18:6445-6445(1990).
[2]
SEQUENCE FROM N.A., AND SEQUENCE OF 65-74; 118-127 AND 165-174.
RN P24268;
RC TISSUE=Liver;
MEDLINE=91354249; PubMed=1883350;
RA Fujita H., Tanaka Y., Noguchi Y., Kono A., Himeno M., Kato K.;
RT "Isolation and sequencing of a cDNA clone encoding rat liver
lysosomal cathepsin D and the structure of three forms of mature
enzymes.";
RL Biochem. Biophys. Res. Commun. 179:190-196(1991).
[3]
SEQUENCE OF 134-170.
RN P24268;
MEDLINE=89034127; PubMed=3182800;
RA Yonezawa S., Takahashi T., Wang X., Wong R.N.S., Hartsuck J.A.,
RA Tang J.;
RT "Structures at the proteolytic processing region of cathepsin D.";
RL J. Biol. Chem. 263:16504-16511(1988).
CC -!- FUNCTION: Acid protease active in intracellular protein breakdown.
CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
CC chain of insulin.
CC -!- SUBUNIT: OCCURS AS A MIXTURE OF BOTH A SINGLE CHAIN FORM AND TWO
CC TYPES OF TWO CHAIN (LIGHT AND HEAVY) FORMS.
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54467; CAA38349.1; .
CC PIR; S13111; KHRTD.
CC PIR; JQ1177; JQ1177.
CC HSP; P07339; ILYB.
CC MEROPS; A01.009; -.
CC InterPro; IPR001461; AspproteaseA1.
CC InterPro; IPR001969; Aspprotease_site.
CC Pfam; PF00026; asp. 1.
CC PRINTS; PR00792; PEPSIN.
CC PROSITE; PS00141; ASP_PROTEASE; 2.

```

```

KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Signal;
Lysosome.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 64 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 65 407 CATHEPSIN D.
FT CHAIN 65 164 CATHEPSIN D 12 KDA LIGHT CHAIN.
FT CHAIN 165 407 CATHEPSIN D 30 KDA HEAVY CHAIN.
FT CHAIN 65 117 CATHEPSIN D 9 KDA LIGHT CHAIN.
FT CHAIN 118 407 CATHEPSIN D 34 KDA HEAVY CHAIN.
FT ACT_SITE 97 97 BY SIMILARITY.
FT ACT_SITE 290 290 BY SIMILARITY.
FT DISULFID 91 160 BY SIMILARITY.
FT DISULFID 110 117 BY SIMILARITY.
FT DISULFID 281 285 BY SIMILARITY.
FT DISULFID 324 361 BY SIMILARITY.
FT CARBOHYD 134 134 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 15 15 D -> A (IN REF. 2).
FT CONFLICT 163 163 D -> T (IN REF. 3).
FT CONFLICT 205 205 K -> N (IN REF. 2).
FT CONFLICT 262 262 K -> N (IN REF. 2).
SQ SEQUENCE 407 AA; 44680 MW; C423AD4104D95F84 CRC64;

Query Match 11.1%; Score 297; DB 1; Length 407;
Best Local Similarity 26.1%; Pred. No. 2.3e-16;
Matches 118; Conservative 76; Mismatches 170; Indels 88; Gaps 20;

Qy 6 PWLLWAGVGLPAHGTHGIRPLR-----SGLGGA--PLGURLPRETDEEPEP 54
Db 4 PGVLLLI-LGLDASSAL-IRPLRKFTSIRMTVEVGSGVEDLILKPTIKYMOSSP 61

Qy 55 GRGSEFVEMVDNLGRKSGQGYVEMTVGSPQTLNLILVDTGSSNFAVGAAPHPL----- 109
Db 62 RTKEPVSELLKNYLDQ---YYEIGIGTPQCFTVVFDTGSSNLWVPSIHCILLDIACW 118

Qy 110 -HRYQRLSSYTRDLRKGVVYPTQCKWEGELGTDLVSPHGPNTVTVRANIAAITESDK 168
Db 119 VHHKYNDSKSTYVKNGTSEDIHYGSGSLSGYLSQDTVSP-----CKSDLGGIKVERQ 172

Qy 169 FF-----INGSNWEGILGLAYAEIARPDSDLEPFFDSLKQTHV-PNLFSLHLCG 217
Db 173 IFGEATKQPGVFIAAKFDGILGNGYFIS--VNVKLPVFDNLKMKQLVKNIFSFYL-- 228

Qy 218 AGFPLAQSEVLASVGGSMIIIGIDHSLYTGLSWYTPIRREWYVEIIVRVEINGDQLKMD 277
Db 229 -----NRDPTGQPGGELMGGTDSRYVHCELSYLVNTRKAYQVHMDQLEV-GSELT- 280

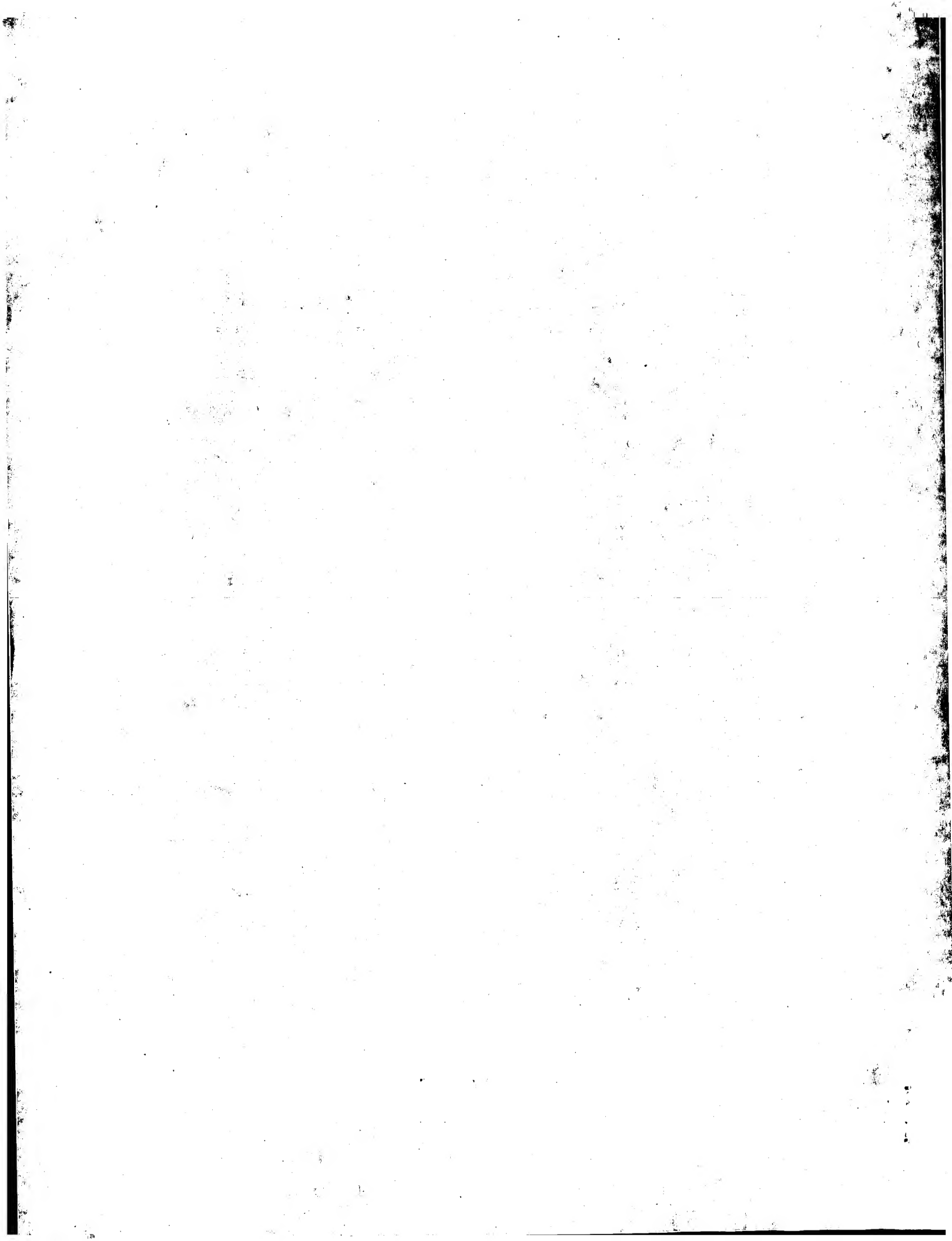
Qy 278 CKENYDKSIVDSGTTNLRPKVFEAAVKSIAASSTKEKFPDGFGLGDLV-CWQAGTT 336
Db 281 CK--GCEAIVDTGTSLLVGPVDEVKELQKAIGAVPLIQ-----GEYNIPEKRVSS- 329

Qy 337 PMNIFPVISLYLMEVNTQSFRTILPQOYLRPVEDVATSDCCYKFAIS-----Q 387
Db 330 ----LPIITFKLGQ-----NYELIPEKYLKVSQAGKT-----ICLSGFMGMDIPPP 373

Qy 388 SSTGTVMGAVIMEGFYVVDRAKRKRGFAVSA 419
Db 374 SGPLWILGDVFIGCYTVTFDREYNRVGFAKAA 405

```

Search completed: June 5, 2003, 16:12:44
Job time : 24 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:11:35 ; Search time 43 seconds
(without alignments)
1120.078 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMLMGAGVLPAAH.....CLRCLRQHQHDDFADDISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Aligned: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2656	99.7	501	2 A59090	aspartic proteinase
2	324	12.2	384	2 JC7574	pepsinogen A - Afr
3	313.5	11.8	385	2 JC7575	pepsinogen A - bul
4	309	11.6	387	2 B38302	pepsin (EC 3.4.23)
5	307.5	11.5	383	2 JC7573	pepsinogen C - Afr
6	307.5	11.5	388	2 S19682	pepsin A (EC 3.4.2)
7	305	11.4	384	2 A39314	gastricsin (EC 3.4)
8	304	11.4	382	1 PECH	pepsin A (EC 3.4.2)
9	301.5	11.3	383	2 A41443	pepsin (EC 3.4.23)
10	301.5	11.3	396	2 A34401	cathepsin E (EC 3)
11	300.5	11.3	412	1 KHHUD	cathepsin D (EC 3)
12	299	11.2	387	2 C38302	pepsin (EC 3.4.23)
13	298	11.2	387	2 D38302	pepsin (EC 3.4.23)
14	297	11.1	407	1 KHRTD	cathepsin D (EC 3)
15	295	11.1	391	2 A43356	cathepsin E (EC 3)
16	294.5	11.1	388	1 S19684	pepsin A (EC 3.4.2)
17	290.5	10.9	380	2 I47176	chymosin (EC 3.4.2)
18	289	10.8	387	2 E38302	pepsin (EC 3.4.23)
19	288.5	10.8	388	1 PEHU	pepsin A (EC 3.4.2)
20	288.5	10.8	388	1 PEMQAR	pepsin A (EC 3.4.2)
21	288.5	10.8	388	2 A30142	pepsin A (EC 3.4.2)
22	288.5	10.8	398	2 S66455	cathepsin E (EC 3)
23	288	10.8	389	2 JE0371	pepsin C (EC 3.4.2)
24	287.5	10.8	444	2 T24204	hypothetical prote
25	287	10.8	410	1 RHMSD	cathepsin D (EC 3)
26	286.5	10.8	388	1 PEMQAJ	pepsin A (EC 3.4.2)
27	286	10.7	398	2 I51195	cathepsin D (EC 3)
28	285.5	10.7	388	2 B30142	pepsin A (EC 3.4.2)
29	284.5	10.7	381	1 CMSHB	chymosin (EC 3.4.2)

ALIGNMENTS

RESULT 1

A59090
aspartic proteinase (EC 3.4.23.-) BACE precursor - human
N:Alternate names: beta-secretase; beta-site APP cleaving enzyme
C:Species: Homo sapiens (man)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000
C:Accession: A59090
R:Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Treplo
M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro
Science 286, 735-741, 1999
A:Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran
A:Reference number: A59090; MUID:2002972; PMID:10531052
A:Note: submitted to GenBank, September 1999
A:Accession: A59090
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-501 <VAS>
A:Cross-references: GB:AF190725; NID:G6118538; PIDN:AAF04142.1; PID:G6118539
C:Genetics:
A:Gene: BACE
C:Superfamily: beta-secretase
C:Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-45/Domain: propeptide #status predicted <PRO>
F:46-501/Product: acid proteinase BACE #status predicted <MAT>
F:461-477/Domain: transmembrane #status predicted <TRN>
F:93-289/Active site: Asp #status predicted
F:153,172,223,354/Binding site: carboxylate (Asn) (covalent) #status predicted
F:330-380/Disulfide bonds: #status predicted

Query Match	99.7%	Score 2656;	DB 2;	Length 501;
Best Local Similarity	99.8%	Pred. No. 7e-206;		
Matches 500;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MAQALPWLMLMGAGVLPAAHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGF	60	
Db	1	MAQALPWLMLMGAGVLPAAHGTQHGIRLPLRSLGGLGAPLGLRLPRETDEEPEEGRGSGF	60	
Qy	61	VEYVDNLRGKSGQGYVEMTVGSPQTINILVDTGSSNFVGAAPHPFLHRYQRQLSST	120	
Db	61	VEYVDNLRGKSGQGYVEMTVGSPQTINILVDTGSSNFVGAAPHPFLHRYQRQLSST	120	
Qy	121	YRDLRKGYVPYPTGKWEGLGTDLVSTPHGPNVTVRANITAAITDESKFFINGNWEGIL	180	
Db	121	YRDLRKGYVPYPTGKWEGLGTDLVSTPHGPNVTVRANITAAITDESKFFINGNWEGIL	180	
Qy	181	GLAYAEIARPDSDLPEPFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI	240	
Db	181	GLAYAEIARPDSDLPEPFDSLVKQTHVNPFLSLHLCGAGFPLNQSEVLASVGGSMIIGGI	240	
Qy	241	DHSLYTSGLWYPTIRREWYVEVITVRVEINGQDLKMDCKEYNDKSVDSGTTNLRPLPKK	300	
Db	241	DHSLYTSGLWYPTIRREWYVEVITVRVEINGQDLKMDCKEYNDKSVDSGTTNLRPLPKK	300	

pepsin A (EC 3.4.2
pepsinogen A - com
chymosin (EC 3.4.2
cathepsin E (EC 3.
aspartic proteinase
gastricsin (EC 3.4
prochymosin - comm
pepsin (EC 3.4.23.
cathepsin D (EC 3.
aspartic proteinase
pepsinogen C - com
pepsinogen (EC 3.4
saccharopepsin (EC
gastricsin (EC 3.4
pepsin A (EC 3.4.2
candidapepsin (EC

Species: *Oryctolagus cuniculus* (domestic rabbit)
#text change 23-Feb-1997

RESULT 5
 JC7573
 pepsinogen C - African clawed frog
 Alternate names: progastricsin
 Species: *Xenopus laevis* (African clawed frog)
 Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
 C:Accession: JC7573; PC7118
 R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
 J. Biochem. 129, 147-153, 2001
 A:Title: Amphibian pepsinogens: Purification and characterization of *Xenopus* pepsinogens
 A:Reference number: JC7573; MUID:21064922; PMID:11134969
 A:Contents: Stomach
 A:Accession: JC7573
 A:Molecule type: mRNA
 A:Residues: 1-383 <IKU>
 A:Cross-references: DDBJ:AB045379
 A:Accession: PC7118
 A:Molecule type: protein
 A:Residues: 17-68 <IK2>
 C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like
 C:Genetics:
 A:Gene: PgC
 C:Superfamily: pepsin
 C:Keywords: stomach; zymogen

 Query Match 11.5%; Score 307.5; DB 2; Length 383;
 Best Local Similarity 25.9%; Pred. No. 6.6e-17;
 Matches 112; Conservative 64; Mismatches 139; Indels 117; Gaps 19;
 Oy 23 QHGIRPLRSLGGLGPIGLRLPRETDEPEPEGRGRSFVEMVDNLRKSGQGYVVENTVG 82

```

A:Molecule type: mRNA
A:Residues: 1-388 <GAG>
A:Cross-references: EMBL:X59753; NID:g38070; PIDN:CAA42425.1; PID:g38071
A:Note: parts of sequence, including amino ends of pepsinogen and activation intermed
C:Comment: This is a minor component of pepsin at all post-partum stages.
C:Comment: Although two-step activation is observed, activation is predominantly a o
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; gastric juice; hydrolase; phosphoprotein; protein di
F:1-15/Domain: signal sequence #status predicted <Sig>
F:16-388/Product: pepsinogen A 4 #status experimental <PPT>
F:16-62/Domain: activation peptide #status experimental <APT>
F:63-388/Product: pepsin A 4 #status experimental <ENZ>
F:38-39/Cleavage site: Leu-Lys (pepsin) #status experimental
F:62-63/Cleavage site: Leu-Ile (pepsin) #status experimental
F:94,277/Active site: Asp #status predicted
F:107-112,268-272,311-344/Disulfide bonds: #status predicted
F:130/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match          11.5%;   Score 307.5;   DB 1;   Length 388;
Best Local Similarity 27.6%;   Pred. No. 6.8e-17;
Matches 108;   Conservative 65;   Mismatches 135;   Indels 83;   Gaps 17;

Qy      44  PRETDEPEEPGRGGSFEMVDNLGRKSGQGYVYVMTGSPQPTLNILVDGSSNFAYGA 103
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      60  PTLIDPEPLE-----NYLDV-----EYFGTIGIGTPAQNTVYVFDTGSSNLVW-- 102

Qy      104  APHPFL-----HRYVYQRLSTYRDLRKGYVVPYTOCKWEGELGTDLVSIHPGNVTV 156
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      103  -PSWVCYSLACDHNLNFPQDSSYRATSKTSYISITGYGTSMTGILGYDTYKVV--GGISD 158

```

157 157 AATAAATESDK-PFINGSNWEGILGLAYAEIARPDSDLEPPFDSLVKQTHV-PNLFSUJH 214
159 159 TNOIFGLSETEPEFFLYAFDGLGLAYPSIS--SSGATPVFDNIWNRQVLSODLFSVY 216
215 215 LCGAGFLPNQSEVLASVGGSMITGGIDHSILYTGSLWYTPIRREWEYVEIIVRVEINGODL 274
217 217 LSAD-----DOS-----GSVVIFGGIDSSYITGSLNWPVSVGYWQISVDSITMGKTI 266
275 275 --KMDCKEYNDKSIDVSGTTLNLRPKKVEAAVKSTKAASSTKPKPDGFWLGEQLV-CW 331
267 267 ACAGC-----QAIVDTGSLTGTSPANTQSDIGASNSD-----GEMVVS 312
332 332 QAGTTPMNPVVISLYLMGEVNTSFRITILPOQY-LRPVEDVATSDDCYK-----FAI 385
313 313 AISSLPDIVF-----TINGVQYPLPPSAIYLOSQSGCTSGFQGMVDP 354
386 386 SOSSTGTVMGAVIMEGFVYVDFDRKRIGFA 416
355 355 TESGELWILGDVFIQYFTVDFDRANNQVGLA 385

RESULT 7
A39314
gastricin (EC 3.4.23.3) precursor - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999
C:Accession: A39314
R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Mikai, K.; Kurokawa, K.; Ito, H.; Kageya
J. Biol. Chem. 266, 22436-22443, 1991
A:Title: Purification, characterization, and amino acid sequences of pepsinogens and pep
A:Reference number: A39314; MUID:92042186; PMID:1939266
A:Accession: A39314
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <XAK>
A:Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 11.4%; Score 305; DB 2; Length 384;
Best Local Similarity 24.5%; Pred. No. 1.1e-16;
Matches 105; Conservative 66; Mismatches 146; Indels 112; Gaps 17;

QY 24 HGIRLPLRSLGAPLGLRLPRETDEPEPEGR--GSFVEMVDNLRGKSGOGYVEMTV 81
DB 35 HGIRAPV-----VDPATKYNNFATAFEPLANVMDMSYVGEISI 73

QY 82 GSPPTNLILVDGTGSSNFAVGAAPHPFL-----HRYQROLSTYRDLRGVVPYQ 134
DB 74 GTPPQNLVLDFTGSSNLW--PSTYCQSQACTNHPQNPQSSSYSSNQOQFSLQYGT 130

QY 135 GKVEGELGTLVSIHPGPNVTRANTA-----AITESDKFFLNGSNWEGILGLAYAE 186
DB 131 GSUTGILGIDYTVQI-----ONIAISQBEGLSVTEPCNTFYV-AQFDGILGLAYS 180

QY 187 IARPDSDLEPPFDSLVKQTHV-PNLFSLHLCGAGFLPNQSEVLASVGGSMITGGIDHSILY 245
DB 181 IA--EGGATVMQGMIOQNLINPLFAFLSGOONSQ-----GGEVAFEGVDQNY 230

QY 246 TGSWLYTPIRREWEYVEIIVRVEINGOD--LMDCKEYNDKSIDVSGTTLNLRPKKVF 302
DB 231 SGQIYWPVTSYQYIGQGFVNGOATCWCSCQ-----QGVIDTGLTLLTAQSYF 284

QY 303 EAAVKSTKAASSTKPKPDGFWLGEQLV-CWQAGTTPMNPVPI-----SLYLMGEVTV 353
DB 285 SSLMQSICAQDQ-----GQYAVSCSNISQSLPTISFTISGVFPLPPSAIYVLOQNS 336

QY 354 NQ-----SFRITILPOQYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGFVYVDFDR 410
DB 337 GYCTIGIMPTVLPSONQPL-----WILGDVFLRQYYSVYDLGN 375

QY 411 KRIGFAVSA 419

376 376 NOVGFAAAA 384

RESULT 8
PECH
pepsin A (EC 3.4.23.1) precursor - chicken
N:Alternate names: pepsinogen A
C:Species: Gallus gallus (chicken)
C:Date: 18-Apr-1984 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C:Accession: JE0370; A00984
R:Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem. Biophys. Res. Commun. 250, 420-424, 1998
A:Title: Analysis of temporal expression pattern and cis-regulatory sequences of chic
A:Reference number: JE0370; MUID:98440813; PMID:9753645
A:Accession: JE0370
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-382 <SAK>
A:Cross-references: GB:AB025281; NID:g4589837; PIDN:BAA76891.1; PID:g4589838
R:Baudys, M.; Kostka, V.
Eur. J. Biochem. 136, 89-99, 1983
A:Title: Covalent structure of chicken pepsinogen.
A:Reference number: A00984; MUID:84004412; PMID:6617663
A:Accession: A00984
A:Molecule type: protein
A:Residues: 16-87, 'S', 89-382 <BAU>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; gastric juice; glycoprotein; hydrolase; protein dige
F:16-57/Domain: activation peptide status experimental <APT>
F:58-382/Product: pepsin A #status predicted <MAT>
F:92,275/Active site: Asp #status predicted
F:105-110,266-270,305-338/Disulfide bonds: #status experimental
F:128/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 11.4%; Score 304; DB 1; Length 382;
Best Local Similarity 24.0%; Pred. No. 1.3e-16;
Matches 88; Conservative 69; Mismatches 125; Indels 84; Gaps 13;

QY 75 YVEMTVGSPPTNLILVDGTGSSNFAVGAAPHPFL-----HRYQROLSTYRDLRGK 127
DB 74 YGFTISIGTPQDFTVIFDTGSSNLW--PSTYCKSSACSNHKKRDPKSSYVSTNET 130

QY 128 VVYPYTOGKWEGLGTLVSIHPGPNVTRANTAIAATESDK-PFINGSNWEGILGLAYAE 186
DB 131 VYIATGSGMSGLIGYDTAV---SSIDVQNLQIFGLSETEPGSFYFFYCNFDGLGLAFPS 187

QY 187 IARPDSDLEPPFDSLVKQTHV-PNLFSLHLCGAGFLPNQSEVLASVGGSMITGGIDHSILY 245
DB 188 IS--SSGATPVFDNMMSQHLVAQDLFSVLSKDG-----ETGSFVLFGGIDPNYT 235

QY 246 TGSWLYTPIRREWEYVEIIVRVEINGODL--MDCKEYNDKSIDVSGTTLNLRPKKVF 303
DB 236 TKGIYWPVLSAETWQITMDRVTVGNKYVACFFTC-----QAIVDTGSLVMPQAYN 289

QY 304 AAVKSIAASSTE-----KFPDGFWLGEQLVCWQAGTTPMNPVVISLYLMGEVNTQS 356
DB 290 RIIRDGLVSSDGEISCDISKLPD-----VTFHNGHA----- 322

QY 357 FRITILPOQYLRPVEDVATSDDCYKFAISOSSTGT-----VMGAVIMEGFVYVDFDR 410
DB 323 -----FTIPASAVVLNEDGSCMLGFENMGPTTELGEOWILGDVFIREYVIFDRAN 373

QY 411 KRIGFA 416
DB 374 NKVGLS 379

RESULT 9
A41443
pepsin (EC 3.4.23.-) precursor, embryonic - chicken
C:Species: Gallus gallus (chicken)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000

R: Hill, J.; Montgomery, D.S.; Kay, J.
FEBS Lett. 326, 101-104, 1993
A: Title: Human cathepsin E produced in E. coli.
A: Reference number: S34467; MUID: 93314762; PMID: 8325357
A: Accession: S34467
A: Status: preliminary
A: Molecule type: protein
A: Residues: 57-60, 62-81 <HIL>
R: Athauda, S.B.P.; Matsuzaki, O.; Kageyama, T.; Takahashi, K.
Biochem. Biophys. Res. Commun. 168, 878-885, 1990
A: Title: Structural evidence for two isozymic forms and the carbohydrate attachment s
A: Reference number: A34643; MUID: 90241267; PMID: 2334440
A: Accession: A34643
A: Status: preliminary
A: Molecule type: protein
A: Residues: 54-58, 'XXX', 62-64, 'M', 66-89, 'X', 91-95 <ATH>
A: Accession: B34643
A: Status: preliminary
A: Molecule type: protein
A: Residues: 54-59, 'X', 61-68 <AT2>
C: Genetics:
A: Gene: GDB: CTSE
A: Cross-references: GDB: 119821; OMIM: 116890
A: Map position: lq31-lq31
C: Superfamily: pepsin
C: Keywords: aspartic proteinase; blocked amino end; hydrolase; zymogen
F: 1-17/Domain: signal sequence #status predicted <SIG>
F: 18-53/Domain: activation peptide #status predicted <PRO>
F: 54-396/Product: cathepsin E #status predicted <NAT>
F: 18/Modified site: blocked amino end (Gln) (in mature form) (probably pyrrolidone ca
F: 96, 281/Active site: Asp #status predicted

Query Match 11.3%; Score 301.5; DB 2; Length 396;
Best Local Similarity 25.8%; Pred. No. 2.1e-16;
Matches 100; Conservative 68; Mismatches 144; Indels 75; Gaps 16;

Qy 48 DEPEPEPGRGSEVEMVDNLGRKSGGYVEMTVGSPQTLNLIVDTGSSNFAVGA--- 103
Db 63 DQSAKEP---LNLVLD-----NEYFGTISIGSPQNFTVIFTDGSNLWVPSVYCT 110
Qy 104 APHFLHRYQROLSTYRDLRKGIVVPYTOGWKEGELGTLVSIOPHPNVTVVRANIAI 163
Db 111 SPACKTHSRFPQSQSTYSQPGQSFSTQYGTGSLGIIGADQVSV-EGLTVVGGQFGEV 169
Qy 164 TESDKFFINGSNWEGILGLAYAIARPDDSLPEFDFSLVKQTHVP-NLFSHLCLGAGFPL 222
Db 170 TEGQTFVD-AEFDGILGLGYPSLA--VGGTVPFVDFNMMAQNLVDLPMSVYM----- 219
Qy 223 NQSEVLASVGSIMIIGDHSLSYTGSLWYTPIRREWYEVIIIVRVEINGDLKMDCKEYN 282
Db 220 -SSNPEGGAGSELIFGQYDHSFSGSLNVVPVTKQAYWQIALDNIQVG--TWMFCE-- 274
Qy 283 YDKSIVDSGTNLRLPKVPFAAVKSIKAASSTEFKPDGFWLGEQLVCWQAGTTPWNIFP 342
Db 275 GCAQIVDTGTSLTIGSPDKIKQLQNAIGAAP-----VDGEYAVE-----CANLNVMP 321
Qy 343 VISLYLMGEVTVNOSFRITLPOQYLRPEVDVATSDQCKFAISQSTG----- 391
Db 322 DVFTTNG-----VPYTLSPATY--TLDFVDGMQFC-----SSGFQGLDIHPAG 365
Qy 392 --TVMGAVIMEGFYVDFRARKRIGFA 416
Db 366 PLWILGDVFIRQFYSVDFRCNNRVGLA 392

RESULT 11.
KHUUD
N: Alternate names: preprocathepsin D
C: Species: Homo sapiens (man)
C: Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000
C: Accession: A25771; S30749; PC20666; I59236; I57716
C: Faust, P.L.; Kornfeld, S.; Chirwin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985
A:Title: Cloning and sequence analysis of cDNA for human cathepsin D.
A:Reference number: A25771; MUID:85270436; PMID:3927292
A:Accession: A25771
A:Molecule type: mRNA
A:Residues: 1-412 <FAD>
A:Cross-references: EMBL:M11233; NID:g181179; PIDN:AAB59529.1; PID:g181180
R:Westley, B.R.; May, F.E.B.
Nucleic Acids Res. 15, 3773-3786, 1987
A:Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast
A:Reference number: S30743; MUID:87231068; PMID:3588310
A:Accession: S30749
A:Molecule type: mRNA
A:Residues: 1-412 <NES>
A:Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678
R:May, F.E.B.; Smith, D.J.; Westley, B.R.
Gene 134, 277-282, 1993
A:Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated
A:Reference number: PC2066; MUID:94085791; PMID:8262386
A:Accession: PC2066
A:Molecule type: DNA
A:Residues: 1-23 <MAY>
A:Cross-references: GB:L12980; NID:g291930; PIDN:AAA16314.1; PID:g455429
A:Experimental source: MCF-7 cell
R:Cavaillès, V.; Augereau, P.; Rochefort, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993
A:Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only
A:Reference number: I59236; MUID:93126342; PMID:8419924
A:Accession: I59236
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV>
A:Cross-references: GB:S5257; NID:g263124; PIDN:AAD13868.1; PID:g4261568
R:Augereau, P.; Miralles, F.; Cavaillès, V.; Gaudelet, C.; Parker, M.; Rochefort, H.
Mol. Endocrinol. 8, 693-703, 1994
A:Title: Characterization of the proximal estrogen-responsive element of human cathepsin
A:Reference number: I57716; MUID:95021301; PMID:7935485
A:Accession: I57716
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <CAV>
A:Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51839; PDB:1LYA
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A:Reference number: A51840; PDB:1LYB
A:Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues 65-
R:Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.; Col
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A:Title: Crystal structures of native and inhibited forms of human cathepsin D: implicat
A:Reference number: A48229; MUID:93342076; PMID:8393577
A:Contents: annotation; X-ray crystallography, 2.5 angstroms
C:Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C:Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolytical
C:Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is bound
C:Genetics:
A:Gene: GDB:CTSD
A:Cross-references: GDB:120512; OMIM:116840
A:Map position: l1p15.5-l1p15.5
C:Function:
A:Description: limited specificity endopeptidase
A:Pathway: intracellular protein degradation
C:Superfamily: pepsin
A:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-64/Domain: propeptide #status predicted <PRO>
F:65-162/169-410/product: cathepsin D #status experimental
F:267-329-356/Region: phosphotransferase recognition
F:91-160-110-117-286-290-329-366/Disulfide bonds: #status experimental
F:97-295/Active site: Asp #status experimental

F:134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 11.3%; Score 300.5; DB 1; Length 412;
Best Local Similarity 26.9%; Pred. No. 2.7e-16;
Matches 123; Conservative 68; Mismatches 170; Indels 97; Gaps 21;
QY 5 LPWLLWMGAGVLPAGHTGIRLPLR-----SGLGGAPLGL-----RLP 44
DQ 7 LPALCLLAA---PASAL---VRPLHKFTSIRRTMSEVGGVEDLIARGPVSKYQAVP 60
QY 45 RETDEPEPEPRGRGSFVMDNLRGSCQGYVEMVGPPTQLNLLVDTGSSNFAVGA 104
DQ 61 AVTE-----GPIPEVLKNYMDAQ---YGEIGIGTTPPQCTVVFVDTGSSNLWPSI 108
QY 105 PHPEL-----HRYORQLSSTYRDLRKGVVYPTQKWEGLGTDLVSP----- 149
DQ 109 HCKLLDTACWILHHKYNDSKSTVYKNGTSPDIHYGSGLSGLYSDTVSPQCSASSA 168
QY 150 HGPVNTVRANIAAITESSDKFFINGSNWEGILGLAYAIARPPDSLEFFDSLVKQTHV-P 208
DQ 169 LGGVKVERQVGEATKQGITFAAKFDGLGMAYPRI--VNNVLPVFDNMQQKLVDO 226
QY 209 NLFSLHLCGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLWYTPIRREWYEVIIVRVE 268
DQ 227 NIFSFL-----SRDPAQPGGELMGLGTDKSKYKGLSYLNTRKAYQWVHLDQVE 278
QY 269 I-NGQDLKMDCKEYNDKSIYDSTGTTNLRPKPKVFEAAVKSIAASSTKPPGDFWLGEO 327
DQ 279 VASGLTL---CKE--GCEAIVDTCTSLMVGPDVRELOKRAIGAVPLIQ-----GEY 325
QY 328 LV-CWQACTTWNIFPVLSLYLMGEVYNQSEFRTILPQOYLRPVEDVATSDQDCYKFAIS 386
DQ 326 MIPCKKST-----LPAITLKLGG---KGYKLS--PEDYTLKVSQAGKTL--CLSGFMG 372
QY 387 Q-----STGVVMGAVIMEGFVYVVDRAKRIGFAVSA 419
DQ 373 MDIPPPSGLWILGDVFTGRTYTVFDRDNNRVGFAEA 410
RESULT 12
C38302
pepsin (EC 3.4.23.-) II-2/3 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Feb-1997
C:Accession: C38302
R:Kageyama, T.; Tanabe, K.; Koizumi, O.
J. Biol. Chem. 265, 17031-17038, 1990
A:Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nu
A:Reference number: A38302; MUID:91009127; PMID:2129536
A:Accession: C38302
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-387 <KAG>
A:Cross-references: GB:J05638
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion
Query Match 11.2%; Score 299; DB 2; Length 387;
Best Local Similarity 26.9%; Pred. No. 3.3e-16;
Matches 97; Conservative 64; Mismatches 134; Indels 66; Gaps 13;
QY 75 YVEMTVGSPPTQLNLLVDTGSSNFAVGAAPHPF-----LHRYORQLSSTYRDLRKG 127
DQ 75 YFGTISIGTPPQDFTVFDTGSSNLW---PSTYCSLLACALHKKRFPEDSDTQGTST 131
QY 128 VYVPTQKWEGLGTDLYSIPHPGNVTVRANIAAITESSDKFFINGSNWEGILGLAYAI 187
DQ 132 LSTYTGSGMTGLGYDTRKVGSIEDTNQIFGLSKTEPSLTFE--APDGILGLAYPSI 189
QY 188 ARPDLSLEPPFDSLVKQTHV--PNLSLHLCGAGFPLNQSEVLASVGSMTIGGIDHSLY 246
DQ 190 SSSDAT--PVFDNNWNEGLVSDQLFVYLSSDD-----EKGLVMFGGIDSYT 237

C:Accession: S13111; C31918; J01177; PQ0222
R:Birch, N.P.; Loh, Y.P.
Nucleic Acids Res. 18, 6445-6446, 1990
A:Title: Cloning, sequence and expression of rat cathepsin D.
A:Reference number: S13111; MUID:91057150; PMID:2243802
A:Accession: S13111
A:Molecule type: mRNA
A:Residues: 1-407 <BIR>
A:Cross-references: EMBL:X54467; NID:g55881; PIDN:CAA38349.1; PID:g55882
R:Yonezawa, S.; Takahashi, T.; Wang, X.; Wong, R.N.S.; Hartsuck, J.A.; Tang, J.
J. Biol. Chem. 263, 16504-16511, 1988
A:Title: Structures at the proteolytic processing region of cathepsin D.
A:Reference number: A92681; MUID:89034127; PMID:3182800
A:Accession: C31918
A:Molecule type: protein
A:Residues: 134-162,'T','I',164-170 <C31918>
R:Fujita, H.; Tanaka, Y.; Noguchi, Y.; Kono, A.; Himeno, M.; Kato, K.
Biochem. Biophys. Res. Commun. 179, 190-196, 1991
A:Title: Isolation and sequencing of a cDNA clone encoding rat liver lysosomal cathepsin D.
A:Reference number: JQ1177; MUID:91354249; PMID:1883350
A:Accession: JQ1177
A:Molecule type: mRNA
A:Residues: 1-14,'A','I',16-204,'N',206-261,'N',263-407 <F0J>
A:Accession: PQ0222
A:Molecule type: protein
A:Residues: 65-74;118-127;165-174 <F0J>
A:Experimental source: liver
C:Comment: Cathepsin D in rat liver lysosome occurs as a mixture of both a single chain and a dimer.
C:Function:
A:Description: limited specificity endopeptidase,
A:Pathway: intracellular protein degradation
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-64/Domain: propeptide #status predicted <PRO>
F:65-407/Product: cathepsin D, 43k single-chain form #status predicted <MAT>
F:65-164/Product: (or 65-165) cathepsin D 12k light chain #status predicted <MA2>
F:65-117/Product: cathepsin D 9K light chain #status predicted <MA4>
F:118-407/Product: cathepsin D 34k heavy chain #status predicted <MA5>
F:165-407/Product: (or 166-407) cathepsin D 30k heavy chain #status predicted <MA3>
F:91-160,110-117,281-285,324-361/Disulfide bonds: #status predicted
F:97,290/Active site: Asp #status predicted
F:134,258/Binding site: carbohydrate (Asn) (covalent) #status predicted

	Query Match	11.1%	Score 297;	DB 1;	Length 407;
	Best Local Similarity	26.1%	Pred. No. 5.le-16;		
	Matches 118;	Conservative 76;	Mismatches 170;	Indels 88;	Gaps 20;
Qy	6	PWLLMNGAGVLPAGHTQHRIPLR-----SGLGGA--PLGLRLPRETDPEEP	54		
Db	4	PGVLLLI-LGLLDASSAL-IRIPLKKFTSIRRTMEVGSGVEDLILKGPIKYSMQSSP	61		
Qy	55	GRGSFVEMVDNLRGKSOGGYIVEMTVGSPQTLNLTDTGSSNFVAAGAHPFL-----	109		
Db	62	RTPKEPVSELLKNLYDAQ---YYGEIGITGPQCFTVFYDGTGSSNLNVSPHCKLLDIAC	118		
Qy	110	-HRYYQRQLSSYTRDLRKGVYPYPTQGWEGELGTDLSIPHGNVTVRANTAAITESDK	168		
Db	119	VHHKYNSDKSSHVVKNKTSTFDIHGSGSLSGVLSQDTSVP-----CKSDUGGIKVEKO	172		
Qy	169	FF-----INGSNWEIGILGAYAEIARPDSDLPPDFDSLVKQTHY-PNLFSHLGCG	217		
Db	173	IFGEATKQPGVVFIAAKFDGILGMGYPFIS--VNKVLPVFDNLMMOKLEKNIFSFL--	228		
Qy	218	AGEPLNQSEVLASVSGSMIIGGIDHSLYTGSLWYTIPREWHYEVIIVRVENQDLKMD	277		
Db	229	-----NRDPTQGPGGELMLGGTDSRYHYHGELSILNVTRKAYQVHMDDLEV-GSELTL-	280		
Qy	278	CKEYNVDKSIIVSDGTTNLRPKKVFEEAAVKSKAAASSTEKFPDPGFWLGEOLV-CWQAGTT	336		
Db	281	CK--GCCEALDTGTSLLVGPVDEVELKQIAGVPLIQ-----GETMIPCERKVS-	329		
Qy	337	PWNIFPVISLYLMGEVNTNSFRITILPOQYLREPVEDVATSQQDCYKFALS-----Q	387		

Fri Jun 6 08:31:00 2003

```

Db      330  ----LPIITFKLGQ-----NYELHPEKVIKVSQAGKT-----ICLSGFMGMDIPPP 373
QY      388  SSGTGVMGAVIMEGFYVVDRAKRGIFAYSA 419
Db      374  SGPLWILGDFIGCYTYVFDREYNRVGFAKAA 405

RESULT 15
A43356
cathepsin E (EC 3.4.23.34) precursor - guinea pig
N:Alternate names: erythrocyte membrane aspartic proteinase; slow-moving proteinase
C:Species: Cavia porcellus (guinea pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A43356
R:Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.; Tanji, M.
J. Biol. Chem. 267, 16450-16459, 1992
A:Title: Gastric procathepsin E and progastricsin from guinea pig. Purification, molecular
A:Reference number: A43356; MUID:92355614; PMID:1644829
A:Accession: A43356
Molecule type: mRNA
Residues: 1-391 <KAG>
A:Cross-references: GB:M88653; NID:g191294; PIDN:AAA37052.1; PID:g191295
A:Note: sequence extracted from NCBI backbone (NCBIN:110763, NCBI:P:110769)
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; membrane protein

Query Match . 11.1%; Score 295; DB 2; Length 391;
Best Local Similarity 26.9%; Pred. No. 7e-16;
Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;

QY      75  YVEMTVGSPPTLNLVDTGSSNFAVGA----AHPFLHRYQRLSTYRDLRKGVV 130
Db      74  YFTISIGSPPTQNTVIFTGSSNLWVPSYCTSPACQTHPVHPSLSSTREVGNFSI 133

QY      131  PYTOGKWEGLGDLVSIHPGPNVTVRANIAAITESDEFFINGSNNEGILGLAYAEIARP 190
Db      134  QYGTGSLTGLIGADQYSV-EGLTVVGOQFGESVQEPKTFVH-ABFDGILGLGYPSLAA- 190

QY      191  DDSLEPFDSLVKOTHPNLFSLHLCAGFPLNQSEVLASVGSMLIGGDHSLYTGSLW 250
Db      191  -GGVTPVFDNNMAQ----NLVALPM----FSVTMSSNPGSGSELTFGGYDFSHFGSLN 241

QY      251  YPIRREMYEVIIVRVEINGQDLKMDCKEYNDKSIDVSGTNNLRLPKKVFEEAAVKS 310
Db      242  WPVTKQAYWQIALDGIQVG--DSVMFCSE--CCQAIVDTGTSITGP----PGKIKQLQ 293

QY      311  AASSTERKFPDGLWGEOLVCWQAGTTPWNIFPVISLYLMGEVTVNQSFRI-----TILPQ 365
Db      294  EALGATVDEGY----SVQC-----ANLNMMLDVT---FINGVPYTLNPTA 333

QY      366  YLRPVEDVATSDDCYKFAISQSSTG-----TVMGAVIMEGFYVVDRAKRG 412
Db      334  Y--TLLDFVDMQVC-----STGFEGLGIOPAGPLWILGDFIRQFYAVFDRGNR 383

QY      413  IGFA 416
Db      384  VGLA 387

```

Search completed: June 5, 2003, 16:17:40
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 16:09:05 ; Search time 84 Seconds
(without alignments)
1228.924 Million cell updates/sec

Title: US-09-869-414A-4
Perfect score: 2664
Sequence: 1 MAQALPWLMMGAGVLPAAH.....CLRLRQOHDFADDISLLK 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2506.5	94.1	476	Q9BYC1	Q9byc1 homo sapien
2	2478.5	93.0	532	Q9ULS1	Q9uls1 homo sapien
3	2398	90.0	457	Q9BYC0	Q9byc0 homo sapien
4	2258.5	84.8	432	Q9BYB9	Q9byb9 homo sapien
5	1407	52.8	266	Q9CUU5	Q9cuu5 mus musculus
6	1155.5	43.4	439	Q9H2V8	Q9h2v8 homo sapien
7	1150	43.2	514	Q9JL18	Q9jll8 mus musculus
8	974.5	36.6	468	Q9NZL2	Q9nzi2 homo sapien
9	969.5	36.4	396	Q9NZL1	Q9nzi1 homo sapien
10	712.5	26.7	213	Q9P0D2	Q9pod2 homo sapien
11	596.5	22.4	255	Q9R1P7	Q9rlp7 mus musculus
12	354.5	13.3	244	Q9WQY9	Q9wqy9 aphrocallis
13	335.5	12.6	391	Q9WKP6	Q9wkp6 drosophila
14	335	12.6	354	Q9GTX7	Q9gtx7 boophilus m
15	319	12.0	384	Q9DEC2	Q9dec2 xenopus lae
16	313.5	11.8	385	Q9DEC4	Q9dec4 rana catesb

17	312.5	11.7	386	6	Q9BGU5	Q9bgu5 bos taurus
18	311	11.7	387	6	Q9GMY8	Q9gmy8 sorex ungui
19	310	11.6	372	5	Q9VLK3	Q9vlk3 drosophila
20	308	11.6	386	6	Q9CMV7	Q9cmv7 rhinolophus
21	307.5	11.5	383	13	Q9DEC3	Q9dec3 xenopus lae
22	307.5	11.5	387	13	Q9DDV5	Q9ddv5 salvelinus
23	307	11.5	387	6	Q9GMY9	Q9gmy9 suncus muri
24	306.5	11.5	383	13	Q9DE45	Q9de45 salvelinus
25	305.5	11.5	376	13	Q9PUR8	Q9pur8 pseudopleur
26	305	11.4	384	13	Q91322	Q91322 rana catesb
27	304	11.4	382	13	Q9PRG9	Q9prg9 gallus gall
28	304	11.4	423	5	Q9VKP7	Q9vkp7 drosophila
29	298.5	11.2	386	6	Q9GMY6	Q9gmy6 canis famil
30	296.5	11.1	396	13	Q93428	Q93428 chionodraco
31	293.5	11.0	381	6	Q9GK11	Q9gk11 camelus dro
32	293	11.0	399	13	Q93458	Q93458 podarcis si
33	290.5	10.9	380	6	Q28950	Q28950 sus scrofa
34	289.5	10.9	399	13	Q9DD89	Q9dd89 brachydanio
35	287.5	10.8	444	5	Q21966	Q21966 caenorhabdi
36	287	10.8	427	5	P91802	P91802 schistosoma
37	286.5	10.8	378	13	Q9PUR9	Q9pur9 pseudopleur
38	284	10.7	389	13	Q9PWK1	Q9pwk1 gallus gall
39	284	10.7	398	13	P87370	P87370 oncorhynch
40	283	10.6	390	6	Q8SQ41	Q8sq41 canis famil
41	283	10.6	422	5	O96906	O96906 onchocerca
42	282.5	10.6	389	13	Q9W643	Q9w643 gallus gall
43	282.5	10.6	390	6	Q9GK10	Q9gk10 camelus dro
44	281.5	10.6	370	6	Q9TTW1	Q9ttw1 bos taurus
45	280.5	10.5	396	13	Q9DEX3	Q9dex3 clupea hare

ALIGNMENTS

RESULT 1

Q9BYC1 PRELIMINARY; PRT; 476 AA.
 AC Q9BYC1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Beta-site APP cleaving enzyme I-476.
 GN BACE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=21408467; PubMed=11516562;
 RA Tanahashi H., Tabira T.;
 RT "Three novel alternatively spliced isoforms of the human beta-site APP
 cleaving enzyme (BACE) and their effect on amyloid beta-peptide
 production.";
 RL Neurosci. Lett. 307:9-12(2001).
 DR EMBL: AB050436; BAB40931.1; -;
 DR HSP; P32329; IYPS.
 DR InterPro: IPR001461; AspproteaseA1.
 DR InterPro: IPR001969; Aspprotease_site.
 DR Pfam: PF00026; asp; 1.
 DR PRINTS: PR00792; PEP SIN.
 DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 476 AA; 52907 MW; 6C8C87F8A953AF66 CRC64;

Query Match 94.1%; Score 2506.5; DB 4; Length 476;
 Best Local Similarity 95.0%; Pred. NO. 6.9e-193;
 Matches 476; Conservative 0; Mismatches 0; Indels 25; Gaps 1;

QY 1 MAQALPWLMMGAGVLPAAHQTGHGIRLPLRSLGAGAPLGLRLPRETDEPEPEGRGSGF 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MAQALPWLMMGAGVLPAAHQTGHGIRLPLRSLGAGAPLGLRLPRETDEPEPEGRGSGF 60

Fri Jun 6 08:31:00 2003

us-09-869-414a-4.rspt

```

QY 61 VEMVDNLKSGGQYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSST 120
Db 61 VEMVDNLKSGGQYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSST 120
QY 121 YRDLKRGVYPYTOGKWEGLGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGTL 180
Db 121 YRDLKRGVYPYTOGKWEGLGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGTL 180
QY 181 GLAYAEIARPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
Db 181 GLAYAEIARPDLSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGI 240
QY 241 DLSLTGSLWTPYIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
Db 241 DLSLTGSLWTPYIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 300
QY 275 DLSLTGSLWTPYIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 360
Db 275 DLSLTGSLWTPYIRREWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKK 360
QY 301 VFEAAVKSIKAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRIT 335
Db 301 VFEAAVKSIKAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRIT 335
QY 335 VFEAAVKSIKAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRIT 420
Db 335 VFEAAVKSIKAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRIT 420
QY 421 ILPOQYLRVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSAC 480
Db 421 ILPOQYLRVEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSAC 480
QY 481 RCLRLRQOHHDFADDISLLK 501
Db 481 RCLRLRQOHHDFADDISLLK 501
QY 456 RCLRLRQOHHDFADDISLLK 476
Db 456 RCLRLRQOHHDFADDISLLK 476

```

```

RESULT 2
Q9ULS1 PRELIMINARY; PRT; 532 AA.
ID Q9ULS1
AC Q9ULS1; 20039618; PubMed=10574461;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE KIAA1149 protein (Fragment).
GN KIAA1149
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
TISSUE=BRIN;
MEDLINE=20039618; PubMed=10574461;
Hirose M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
"Characterization of cDNA clones selected by the GeneMark analysis
from size-fractionated cDNA libraries from human brain.";
DNA Res. 6:329-336(1999).
RL EMBL; AB032975; BAA86463.2;
DR HSP; P56272; 1AM5.
DR MEROPS; A01.004;
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 532 AA; 58720 MW; 98B135D0B5FBD2E8 CRC64;

Query Match 93.0%; Score 2478.5; DB 4; Length 532;
Best Local Similarity 96.1%; Pred No. 1.4e-190;
Matches 473; Conservative 1; Mismatches 15; Indels 3; Gaps 2;

QY 11 WMGAVLIP-AHGTQHGIRLPSLGGAPLGLRLPRETDEPEPGRGSEVMVDNLRG 69
Db 43 WARECCLPATPASTAG--CPCAAAGGAPLGLRLPRETDEPEPGRGSEVMVDNLRG 100

```

```

QY 70 KSGOGYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSSTYRDLKRGVY 129
Db 70 KSGOGYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSSTYRDLKRGVY 129
QY 101 KSGOGYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSSTYRDLKRGVY 160
Db 101 KSGOGYVEMTVGSPPTNLILVDTGSSNFAVGAAPHPFLHRYIQRLSSTYRDLKRGVY 160
QY 130 VPYTQGWEGELGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGILGLAYAEIAR 189
Db 130 VPYTQGWEGELGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGILGLAYAEIAR 189
QY 161 VPYTQGWEGELGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGILGLAYAEIAR 220
Db 161 VPYTQGWEGELGTLVSIHPGPNVTYRANIAAATESDKFFINGSNWEGILGLAYAEIAR 220
QY 190 PDSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSL 249
Db 190 PDSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSL 249
QY 221 PDSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSL 280
Db 221 PDSLEPFDSLVKQTHVFNLSLHLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSL 280
QY 250 WYTPILRRWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKKVFEEAAVKSI 309
Db 250 WYTPILRRWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKKVFEEAAVKSI 309
QY 281 WYTPILRRWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKKVFEEAAVKSI 340
Db 281 WYTPILRRWYEVIIIVRVEINGQDLKMDCKEYNDKSIIVDSGTTNLRPKKVFEEAAVKSI 340
QY 310 KAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRITILPQOYLRLP 369
Db 310 KAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRITILPQOYLRLP 369
QY 341 KAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRITILPQOYLRLP 400
Db 341 KAASSTKEKPPDGFWMGEQLVQWAGTTPWNIPVLSLYLMGEVNTNOSFRITILPQOYLRLP 400
QY 370 VEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSACHVDEPRTA 429
Db 370 VEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSACHVDEPRTA 429
QY 401 VEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSACHVDEPRTA 460
Db 401 VEDVATSDDCYKFAISQSSTGTVMGAVIMEGFYVYVDRARKRIGFAVSACHVDEPRTA 460
QY 430 AVEGPPVTLDMECDGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCWRCLRLRQ 489
Db 430 AVEGPPVTLDMECDGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCWRCLRLRQ 489
QY 461 AVEGPPVTLDMECDGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCWRCLRLRQ 520
Db 461 AVEGPPVTLDMECDGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCWRCLRLRQ 520
QY 490 HDDFADDISLLK 501
Db 490 HDDFADDISLLK 501
QY 521 HDDFADDISLLK 532
Db 521 HDDFADDISLLK 532

RESULT 3
Q9BYC0 PRELIMINARY; PRT; 457 AA.
ID Q9BYC0
AC Q9BYC0; 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Beta-site APP cleaving enzyme I-457 (Beta-site APP cleaving enzyme
type C).
GN BACE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
TISSUE=BRIN;
MEDLINE=21408467; PubMed=11516562;
Tanahashi H., Tabira T.;
"Three novel alternatively spliced isoforms of the human beta-site APP
cleaving enzyme (BACE) and their effect on amyloid beta-peptide
production.";
RL Neurosci. Lett. 307:9-12(2001).
RN [2]
RP SEQUENCE FROM N.A.
TISSUE=EXOCRINE PANCREAS;
Zacchetti D., De Pietri Tonelli D., Schnurbus R.;
"New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
human pancreas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AB050437; BAB40932.1;
DR EMBL; AF338817; AAK38375.1;
DR HSP; P32329; 1YPS.
DR InterPro; IPR001461; AspproteaseA1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 457 AA; 51068 MW; C79AC9A9E85FE7A2 CRC64;

```

Query Match 90.0%; Score 2398; DB 4; Length 457;
 Best Local Similarity 91.0%; Pred. No. 3.4e-184;
 Matches 456; Conservative 0; Mismatches 1; Indels 44; Gaps 1;

QY 1 MAQALPWLILWMAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
 DB 1 MAQALPWLILWMAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
 QY 61 VEMVDNLGRKSGGYVEMTVGSPPTLNLVDTGSSNFAGVGAAPHFLHRYTORQLSST 120
 DB 61 VEMVDNLGRKSGGYVEMTVGSPPTLNLVDTGSSNFAGVGAAPHFLHRYTORQLSST 120
 QY 121 YRDLRGVYVPYTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
 DB 121 YRDLRGVYVPYTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 240
 DB 146 -----PDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 196
 QY 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 300
 DB 197 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 256
 QY 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQCWAGTTPWNIFFPVISLYLMGEVNTNQSFRIT 360
 DB 257 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQCWAGTTPWNIFFPVISLYLMGEVNTNQSFRIT 316
 QY 361 ILPQOYLRPVEDVATSDODCCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 420
 DB 317 ILPQOYLRPVEDVATSDODCCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 376
 QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
 DB 377 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 436
 QY 481 RCLRLRQHQHDDFADDISLLK 501
 DB 437 RCLRLRQHQHDDFADDISLLK 457

RESULT 4
 Q9BYB9 ID Q9BYB9 PRELIMINARY; PRT; 432 AA.
 AC Q9BYB9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE Beta-site APP cleaving enzyme I-432.
 GN BACE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=21408467; PubMed=11516562;
 RA Tanahashi H., Tabira T.;
 RT "Three novel alternatively spliced isoforms of the human beta-site APP
 RT cleaving enzyme (BACE) and their effect on amyloid beta-peptide
 RT production.";
 RL Neurosci. Lett. 307:9-12(2001).
 DR EMBL; AB050438; BAB40933.1; -;
 DR HSP; P32329; IYPS.
 DR InterPro; IPR001461; AsproteaseA1.
 DR InterPro; IPR001969; AsproteaseA1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PREPIN.
 DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 432 AA; 48212 MW; 96FC81EGF0EED01B CRC64;

Query Match 84.8%; Score 2258.5; DB 4; Length 432;

Best Local Similarity 86.2%; Pred. No. 5e-173;
 Matches 432; Conservative 0; Mismatches 0; Indels 69; Gaps 1;

QY 1 MAQALPWLILWMAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
 DB 1 MAQALPWLILWMAGVLPAGHTGIRLPLRSLGGLGAPLGLRLPRETDEPEPGRGSGF 60
 QY 61 VEMVDNLGRKSGGYVEMTVGSPPTLNLVDTGSSNFAGVGAAPHFLHRYTORQLSST 120
 DB 61 VEMVDNLGRKSGGYVEMTVGSPPTLNLVDTGSSNFAGVGAAPHFLHRYTORQLSST 120
 QY 121 YRDLRGVYVPYTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
 DB 121 YRDLRGVYVPYTGKWEGLGDLVSIHPGPNVTVRANTAAITESDKFFINGSNWEGIL 180
 QY 181 GLAYAEIARDDSLPEFFDSLVKQTHVFNLSLHLCGAGPLNQSEVLASVGGSGMIIGGI 240
 DB 146 -----LCGAGPLNQSEVLASVGGSGMIIGGI 171
 QY 241 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 300
 DB 172 DHSLYTGSLSWYTPIRREWYVEIIVRVEINGQDLKMDCKEYNDKSIDVSGTTLNRLPKK 231
 QY 301 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQCWAGTTPWNIFFPVISLYLMGEVNTNQSFRIT 360
 DB 232 VFEAAVKSIIKAASSTKFPDGFGLGEOLVQCWAGTTPWNIFFPVISLYLMGEVNTNQSFRIT 291
 QY 361 ILPQOYLRPVEDVATSDODCCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 420
 DB 292 ILPQOYLRPVEDVATSDODCCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSAC 351
 QY 421 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 480
 DB 352 HVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDSTLMTIAYVMAAICALFMLPLCLMVCOW 411
 QY 481 RCLRLRQHQHDDFADDISLLK 501
 DB 412 RCLRLRQHQHDDFADDISLLK 432

RESULT 5
 Q9CUU5 ID Q9CUU5 PRELIMINARY; PRT; 266 AA.
 AC Q9CUU5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Beta-site APP cleaving enzyme (Fragment).
 GN BACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=BRAIN;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyono-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

61 SKGFDVTYKTYGGSWTGFVGEDLWTIPKGFNTSFLYNIAITIPESNFFLPGIKWNGILGL 120
 183 AYAETARPDSSLEPFPSLSLVKQTHVPNLFSLHLCGAGFPQNSSEVLASVGGSMIIGGDH 242
 121 AYATLAKSSSELTFFPSLVTQANIPNVSQMGAGLPVAGS--GTNGGSLVLGGTGP 177
 243 SLYTGLWYTPIRREWYEVIIIRVEINGODLKMCKEYNDKSIYDVSCTTNLRLPKKVF 302
 178 SLYKGDWYTPIKEWYQIIEILKEIGGSLNDCREYNADKAIVDSGTTLLRLPQKVF 237
 303 EAAVYSIKAASTSEKFPDGFGLGQVLCWQAGTTPWNIPFVSIYSLMLGVTNQSFRTIL 362
 238 DAVEAVARASLIPESDFGFWTGSQACWTNSETPNWSYFPKSIYLRDENSRSFRITIL 297
 363 PQOYLRPVEDVATSDDCYKFAISOSSTGTVMGAVIMEGYVYVDFRARKRIGFASACHV 422
 298 POLYIQPMAGLNT--ECYREGISPTNALVIGATMEGYVDFRAQKRVGFASPCAE 356
 423 HDEFRTAAVEGFPFVILDMEDCGYNIPQDTSTLTIAYVMAAIC--ALFMLPLCLMVCWR. 481
 357 IAGAAYSEISGFPSEFVEDVASNCVPAQSLSEPIILVISALMSVCGAILLVILVLLLPFR 416
 482 CLRCLRQHQHDFADDISLL 500
 417 CQR--RPRDPEVVNDESSL 433
 [1]
 PRELIMINARY; PRT; 514 AA.
 Q9JUL18
 AC Q9JUL18; 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Aspartyl protease 1.
 GN BACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Graniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 SEQUENCE FROM N.A.
 RA Choi D.K., Sugano S., Sakaki Y.;
 RT "Molecular characterization of the mouse Aspl gene, a homolog of the
 human Aspl (Down Syndrome Region aspartyl protease).";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF216310; AAF36599.1; -.
 DR HSSP; P32329; 1YFS.
 DR MEROPS; A01.041; -.
 DR MGD; MGI:1860440; Bace2.
 DR InterPro; IPR001461; AspproteaseA1.
 DR InterPro; IPR001969; Aspprotease_site.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
 KW Protease.
 SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;
 Query Match 43.2%; Score 1150; DB 11; Length 514;
 Best Local Similarity 48.3%; Pred. No. 7e-84;
 Matches 224; Conservative 76; Mismatches 144; Indels 20; Gaps
 14 AGVLPAGTQHGIRLPLRSGLGAGPLGLRLPRETDEEPEEGRGRGSEVMVDNLKGSQ 733
 43 ASAVPGIGTTP--ELPRADGLA--LALEPVRAI-----ANFLAMVDNLGDSGR 86
 74 GYVEMTVGSPQOTNLINILVDTSNFAVCAAPHPLHRYQVRLSSYTRDLRKGVVPT 13
 87 GYVLEMLIGTPPKQVQILVDTSNFAVAGAPHYSYIDYFDESSESYTHSKGSDVTKY 14
 134 QGKWEGLGTDLIVSIPIHGNNVVRANIAITESDKDFINGSNWIGILGLAYAEIARPDSS 15

Db 147 QGWTGFGVDELVTIPKGFNSFLVNIATIFESNFFLPGIKWNGILGLAYAAAKPSS 206
Qy 194 LEPPFDSLKVQTHVPLNFSHLGCGPPLNOSVELASVGGSMIIGDHSLYTGLWYTP 253
Db 207 LETFDSILVAQAQIPDIFSMQCCAGLPVAGS---GTNGGSLVGLGIEPSLYKGDWYTP 263
Qy 254 IRREWYEVIVRVEINGQDLKMDCKEYNDKSTVDSGTNNLRPKVKFEAAVKSIRKAS 313
Db 264 IKREWYQIEILKLEIGQNLNLDREYNADKATVDSGTTLLRLPKQKVFDAVAVARTS 323
Qy 314 STEKFPDGLWELGECVQAGTTPWNIFFVLSLYLMGEVNTNQSFRITLPOQLRPVEDV 373
Db 324 LIPEFGDFTGAQACWTNSETPWAYFPKISIVLRDENASRSFRITLIPOLYQPMGA 383
Qy 374 ATSDQCYKFAISOSSTGTVMGAVIMEGFYVDFRARKRIGFAVSACHVHDEFRTAAVEG 433
Db 384 GFNY-ECYRGISSTNALVIGATVMEGFYVDFRARRVGFVSPCAEIEGTVSEISG 442
434 PFVTLMDCCGYNIPTQDESTLMTIAYMAICALFMLPLCLMV 477
443 PFSTEDIASNCVPAQALNEPTLWISYALMSVCGAILLVLL 486

RESULT 8

Q9NZL2 ID Q9NZL2 PRELIMINARY; PRT; 468 AA.
AC Q9NZL2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase."
RL Cyogenet. Cell Genet. 89:177-184(2000).
DR EMBL; AF188276; AAF35835.1;
DR HSSP; P00797; 2REN.
DR MEROPS; A01.041;
DR InterPro; IPR001461; Asparticase1.
DR InterPro; IPR001969; Asparticase_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPFIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
DR Protease.
SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Query Match 36.68; Score 974.5; DB 4; Length 468;
Best Local Similarity 40.5%; Pred. No. 7.6e-70;
Matches 210; Conservative 76; Mismatches 150; Indels 83; Gaps 10;

Qy 2 AQAALPWLKLM---GAGVLPAGHTQHGIRLPLRSGLG-----GAPL-----GLR 42
Db 7 ALLPLLAQMLRAAPLAPPT-----LPLRVAATNRVAPTPGPTPAERHADGLA 61
Qy 43 LPRETDEPEEPGRGRSFVEMVNLGRKSGGYVEMTVPOTNLILVDGTSSNFAVG 102
Db 62 LALE--PALASPAAGANFLAMVDNLQDSCGRGYLEMLIGTPPKQLIILVDGTSSNFAV 119
Qy 103 AAPHPFLHRYQROLSTYRDLRGVYVYPTQGWEGELGDLVSIHPGNVTVRANTAA 162
Db 120 GTPHSYIDTVFDTSTYRSKGFDTVKYVQGSWTGFGVDELVTIPKGFNTSLVNIAT 179
Qy 163 ITESDKFEFNGSNWEGILGLAYAEIARPDOSLEPFFDSLKVQTHVPLNFSHLGCGAPPL 222
Db 180 IFESNFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLTQANIPNVFSNQMGAGLPV 239

Qy 223 NQSEVLASVGGSMIIGDHSLYTGLWYTPIRREWYEVIVRVEINGQDLKMDCKEYN 282
Db 240 AGS---GTNGGSLVGLGIEPSLYKGDWYTPKEWYQIEILKLEIGQSLNLDREYN 296
Qy 283 YKSIDVSGTTNNLRPKVKFEAAVKSIRKASSTKEKFPDGLWELGECVQAGTTPWNIFF 342
Db 297 ADKAIIVDSGTTLLRLPKQKVFDAVAVARTS----- 329
Qy 343 VISLYLAGEVNTNQSFRITLPOQLRPVEDVATSDQCYKFAISOSSTGTVMGAVIMEGF 402
Db 330 -----YIQPMMGAGLNY-ECYRGISSTNALVIGATVMEGF 365
Qy 403 YVDFRARKRIGFAVSACHVHDEFRTAAVEGFYVDFRARRVGFVSPCAEIEGTVSEISG 462
Db 366 YVDFRARRVGFVSPCAEIEGTVSEISG----- 425
Qy 463 AAIC-ALFMLPLCLMVCMWRCLRQOHDDFADDSLL 500
Db 426 MSVCGAILLVLLVLLLPFRQOR--RPRDPEWVNDSSSL 462

RESULT 9

Q9NZL1 ID Q9NZL1 PRELIMINARY; PRT; 396 AA.
AC Q9NZL1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase."
RL Cyogenet. Cell Genet. 89:177-184(2000).
DR EMBL; AF188277; AAF35836.1;
DR HSSP; P00797; 2REN.
DR MEROPS; A01.041;
DR InterPro; IPR001461; Asparticase1.
DR InterPro; IPR001969; Asparticase_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPFIN.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_2.
DR Protease.
SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match 36.44; Score 969.5; DB 4; Length 396;
Best Local Similarity 49.3%; Pred. No. 1.5e-69;
Matches 200; Conservative 56; Mismatches 111; Indels 39; Gaps 9;

Qy 2 AQAALPWLKLM---GAGVLPAGHTQHGIRLPLRSGLG-----GAPL-----GLR 42
Db 7 ALLPLLAQMLRAAPLAPPT-----LPLRVAATNRVAPTPGPTPAERHADGLA 61
Qy 43 LPRETDEPEEPGRGRSFVEMVNLGRKSGGYVEMTVPOTNLILVDGTSSNFAVG 102
Db 62 LALE--PALASPAAGANFLAMVDNLQDSCGRGYLEMLIGTPPKQLIILVDGTSSNFAV 119
Qy 103 AAPHPFLHRYQROLSTYRDLRGVYVYPTQGWEGELGDLVSIHPGNVTVRANTAA 162
Db 120 GTPHSYIDTVFDTSTYRSKGFDTVKYVQGSWTGFGVDELVTIPKGFNTSLVNIAT 179
Qy 163 ITESDKFEFNGSNWEGILGLAYAEIARPDOSLEPFFDSLKVQTHVPLNFSHLGCGAPPL 222
Db 180 IFESNFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLTQANIPNVFSNQMGAGLPV 239

```

223 NQSEVLASVCGSMIGIDHSYTGSLWYTPIRREWEYEVIIIVRVEINGDQKMDCKEYN 282
240 AGS---GTNGSGSVLGGIEPSYKGDWYTPKEEYQYIEILKLEIGQSLNLDREYN 296
283 YKSDVSGTNTLRPKKVFEEAVKSIKAASSTKEKFPDGFGLGECQVLCWQAGTTTWNIFP 342
297 ADKAIVDSGTTLLRLPKQKVEDAVAVARASLIPEESDGFGLGECQVLCWQAGTTTWNIFP 356
343 VISIYLMGEVNTSFRITILPOOYLPRVEDVATSDQDCYKF-AISO 387
357 KISYLRDENSRSFRITILPOK-LRVLQ-----CLKFPLGSLQ 393

RESULT 10
Q9P0D2 PRELIMINARY; PRT: 213 AA.
AC Q9P0D2;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
HSPCL04 (Fragment).
Homo sapiens (Human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=CORD BLOOD;
RA Zhang Q.H., Ye M., Zhou Y., Wu X.Y., Guan Z.Q., Wang L.,
RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT "Human partial CDS cloned from cd34+ stem cells.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF161367; AAF28927.1;
DR InterPro: IPR001461; AspproteaseA1.
DR Pfam: PF00026; asp. 1.
DR NON_TER 1
FT SEQUENCE 213 AA; 24338 MW; EC9D3FA31CFA835C CRC64;
SQ
Query Match 26.7%; Score 712.5; DB 4; Length 213;
Best Local Similarity 83.5%; Pred. No. 2.6e-49;
Matches 137; Conservative 4; Mismatches 12; Indels 11; Gaps 1;

Qy 238 GGDHSYTGSLWYTPIRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSIDSGTTNLR 297
Db 1 GGDHSYTGSLWYTPIRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSIDSGTTNLR 60
Qy 298 PKKVFEAAVKSIRKAASSTKEKFPDGFGLGECQVLCWQAGTTTWNIFPVISIYLMGEVNTNSF 357
61 PKKVFEAAVKSIRKAASSTKEKFPDGFGLGECQVLCWQAGTTTWNIFPVISIYLMGEVNTNSF 120
358 RITILPOOYLPRVEDVATSDQDCYKF-AISO 401
121 RITILPOOYLPR-----WKMPRPKTTVTCHLTVIHG 153

RESULT 11
Q9R1P7 PRELIMINARY; PRT: 255 AA.
AC Q9R1P7;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Aspartyl protease (Fragment).
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A.
RA Accorino M., Fumagalli P., Taramegli R., Ottolenghi S.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
transmembrane protease.";

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF051150; AA045964.1;
DR MEROPS: A01.041;
DR MGD: MGI:1860440; Bace2.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR PRINTS: PR00792; PPSIN.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
KW Protease.
FT NON_TER 1
SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;
Query Match 22.4%; Score 596.5; DB 11; Length 255;
Best Local Similarity 47.8%; Pred. No. 7e-40;
Matches 109; Conservative 44; Mismatches 74; Indels 1; Gaps 1;

Qy 250 WYTPIRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSIDSGTTNLRPKKVFEEAVKSI 309
Db 1 WYTPIRREWEYEVIIIVRVEINGDQKMDCKEYNVDKSIDSGTTNLRPKKVFEEAVKSI 60
Qy 310 KAASSTKEKFPDGFGLGECQVLCWQAGTTTWNIFPVISIYLMGEVNTNSFRITILPOOYL 369
61 ARTSLIPEESDGFGLGECQVLCWQAGTTTWNIFPVISIYLMGEVNTNSFRITILPOLIOP 120
370 VEDVATSDQDCYKF-AISO 429
121 MMGAGFNY-ECYRFGISSSTNALVIGATVMEGFYVVFDRQRRVGFVSPCAEIEGTIVS 179
430 AVEGPPVTLDMEDCGYNIPQDTDESTLMTIAYVMAICAFMLPLCLMV 477
180 EISGPESTEDIASNCVPAQALNEPILWISYALMSVCGAILLVILL 227

RESULT 12
Q8WQY9 PRELIMINARY; PRT: 244 AA.
AC Q8WQY9;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Aspartate protease (Fragment).
GN APP.
OS Aphrocallistes vastus.
OC Eukaryota; Metazoa; Porifera; Hexactinellida; Hexasterophora;
OC Lysacnoids; Rosellidae; Aphrocallistes.
OX NCBI_TaxID=83887;
[1]
RN RP SEQUENCE FROM N.A.
RA Mueller W.E.G., Mueller I.M., Grebenjuk V.A.;
RT "Urmatazoa: Origin and evolution of the common ancestor of Metazoa.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ304863; CAC83293.1;
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR InterPro: IPR000005; HTHArac.
DR Pfam: PF00026; asp. 1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Protease.
FT NON_TER 244
SQ SEQUENCE 244 AA; 26366 MW; 6536902661E0E4C7 CRC64;
Query Match 13.3%; Score 354.5; DB 5; Length 244;
Best Local Similarity 35.6%; Pred. No. 1.8e-20;
Matches 80; Conservative 41; Mismatches 91; Indels 13; Gaps 6;

Qy 64 VDNLRCKSGQYVEMTVGSPDPTNLIVDTGSSNFVAGAAPHPFLHRY--YQQLSSTY 121
22 VYOLQGPESGYLSVNLGTPPEQEFKVLVDGSSNFVAGAAPHPFLHRY--YQQLSSTY 81
122 RDLRGVYVYPTQGGKEGELGDLVS----TPHGPNVIVRANIAL--TESDKFELGNSW 176
82 RDINSEVGKVIDGWSNGRVEDVAFASDVTTNASKSVKVVYSLIESVSEGFFETSGW 141
```



```

DR InterPro: IPR001461; AspproteaseA1.
DR Pfam: PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
FT NON_TER 1
SQ SEQUENCE 354 AA; 38512 MW; B50D7C638CF27091 CRC64;

Query Match 12.6%; Score 335; DB 5; Length 354;
Best Local Similarity 26.2%; Pred. No. 1.2e-18;
Matches 101; Conservative 75; Mismatches 134; Indels 76; Gaps 15;

QY 64 VDNLRK-----SGGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFL- 109
DB 13 VTEIRGALGDPTPIILTYNNMQFYGIITCTPQSFLLMDTGSSNEWV-----PSTN 67

QY 110 -----HRYQROLSTYRDLRGVVPYTOGKWEGLGTDLVSIIPHGNVTVRANI 160
DB 68 CDOSMACRDRHAKYDSKSTFTKSGRIIRIRVSGGVVGRITSIDNVG--GRATVTOYKF 125

QY 161 AATESDKFFINGSNWEGILGLAYAEIARPDPSLEFFDSLVKQTHVPN-LFSLHLCGAG 219
DB 126 AEMHSDGKLFRAKYDGIFGLAFPSISQ--NNQLPLFDAMVKQGVVROAVESLYL--SK 181

QY 220 PFLNQSEVLASVGGMIIGDIDISLYTGSWYTPIRREWYEVIIIVRVEINGODLKM-DC 278
DB 182 QPSEQN-----GGEYFGGINAQRYTGAHYVPSQAARHWOVMDNINNVQGTTLGVGC 235

QY 279 KEYNDKSIIVDSTNLRPKKVFEEAAVKSIIKAASSTERKFPDGFGLGEOLVCWQAGTTPW 338
DB 236 -----PTVVDSTGTSFLSGP-----SADVELNRVIGATKTAAGY-----FEVNCAT 277

QY 339 NIFPVISLYLMGE---VTNQSFRTILPOQYLRPVEDVATSDQDCYKFAISQSTGT--- 392
DB 278 SSLRPITFNLNGKSFPLQGEAYTIRI-----PLTTGGGQCFTRISESASGTNLW 327

QY 393 VMGAVIMEGYVVDRAKRIGFAVS 418
DB 328 ILGAVETQTYTYVFDRAQNRVGPATA 353

```

RESULT 15

```

Q9DEC2 Q9DEC2 PRELIMINARY; PRT; 384 AA.
AC Q9DEC2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Pepsinogen A.
OS Xenopus laevis (African clawed frog).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
RA Ikuzawa M., Inokuchi T., Kobayashi K., Yasumasu S.;
RT "Molecular Cloning of Pepsinogens in Adult Xenopus laevis and Bullfrog
RT Rana catesbeiana."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
DR EMBL; AB045380; BAB20798.1; -
DR HSSP; P00790; IPSN.
DR MEROPS; A01.001; -.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
KW Aspartyl protease; Hydrolase.
SQ SEQUENCE 384 AA; 41888 MW; 729DD2E7953D9072 CRC64;

```

Query Match 12.0%; Score 319; DB 13; Length 384;
 Best Local Similarity 27.0%; Pred. No. 2.5e-17;

```

Matches 118; Conservative 70; Mismatches 163; Indels 86; Gaps 21;

QY 8 LLLWMGAGVLPAGHQHGIHLPLRLSLGLG-----APLGLRLPRETDEEPEEPGRK----- 57
DB 3 LLLLLGLVVL-----SECVVKVPLRKGESFRNLQRLGL-LGDLKKNYPNPASKYFPTLA 57

QY 58 GSFVEMVDNLRGKSGQGYVEMTVGSPPTLNLIVDTGSSNFAVGAAPHPFL-----H 110
DB 58 QSSAETLQNYMDIE---YGTISICTPPOEFTVIEDTGSANLW---PSVYSSQACSNIH 111

QY 111 RYRQRLSSTYRDLRGVVPYTOGKWEGLGTDLVSIIPHGNVTVRANIAITESDK-F 169
DB 112 NRENFQOQSTFOATNTPVSIQYGTGSMGFLGYDTLQV---CNIQISNMFLSESEPGS 168

QY 170 FINGSNWEGILGLAYAEIARPDPSLEFFDSLVKQTHVP-NLFSHLHLCGAGFPLNQSEVL 228
DB 169 FIYYSPFDGILGLAFPSIA--SSQATPVFDNNWSOGLIPQNLFSVYLSDDG----- 217

QY 229 ASVGGSMIIGTDHSLYTGSLWYTPIRREWYEVIIIVRVEINGODL--KMDCKEYNYDKS 286
DB 218 -QTGSVVLFGGVDSYSGSLNNVPLTAETYWQITLDSVSIQGVIAQSQC-----QA 270

QY 287 IYDSTTNLRPLPKKVFEEAAVKSIIKAASSTERKFPDGFGLGEOLV-CWQAGTTPNIFPVIS 345
DB 271 IYDTGTSMTGPSTPI-ANIQYIGASQDSN-----GQYVINCNNISNMPTIIVF----- 318

QY 346 LYLMGEVNTQSFRTILPOQYL-RPVEDVATSDQDCYK-FAISQSTGT---VMGAVIM 399
DB 319 -----TINGVOYPLSPSAYVRQNOGCGSSGFOAMNLPNTNSGDLWILGDVFI 364

QY 400 EGYVYVVDRAKRIGFA 416
DB 365 RQYFTVFDRAKNYVAIA 381

```

Search completed: June 5, 2003, 16:14:17
 Job time : 87 secs